STIC-Biotech/ChemLib

From: Steadman, David (AU1652)
Sent: Wednesday, May 30, 2001 8:03 AM
To: STIC-Biotech/ChemLib
Subject: 09/526,193 SEQ SEARCH

NAME: David Steadman
AU: 1652
Date: 05/30/01
Room: 10D-04
Mailbox #: 10C-01 M3
Serial #: 09/526,193

Please search the following sequence(s) in commercial databases:

Amino Acid
Amino acids 1-60 of SEQ ID NO:1 against amino acid databases
Amino acids 1-60 of SEQ ID NO:1 against nucleic acid databases

Thank you,
David Steadman

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 31, 2001, 13:04:34 ; Search time 49.05 Seconds
(without alignments)
143.374 Million cell updates/sec

Title: US-09-526-193A-1_COPY_1_60
Perfect score: 334
Sequence: 1 MACWPQLRLLLWKNLTFRRR.....SVRLSYPPYEQHECHFPNKA 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	334	100.0	70	4	Q9NS76	Q9ns76 homo sapien
2	334	100.0	100	4	Q9NP93	Q9np93 homo sapien
3	334	100.0	2261	4	Q9NQV4	Q9nqv4 homo sapien
4	230	68.9	2146	4	Q9NR73	Q9nr73 homo sapien
5	185	55.4	2310	11	O35600	O35600 mus musculu
6	179	53.6	2281	6	O02698	O02698 bos taurus
7	100	29.9	602	5	Q9N403	Q9n403 caenorhabdi
8	78	23.4	1713	5	Q9VRG4	Q9vrg4 drosophila
9	71.5	21.4	707	10	Q9T050	Q9t050 arabidopsis
10	65.5	19.6	307	13	Q9PVW2	Q9pvw2 oryzias lat
11	65.5	19.6	307	13	Q9PVW1	Q9pvw1 oryzias lat
12	61.5	18.4	129	10	Q9ZPB9	Q9zpb9 crawfurdia
13	61.5	18.4	2272	5	Q9VZS7	Q9vzs7 drosophila
14	61	18.3	218	5	Q9NF52	Q9nf52 drosophila
15	60.5	18.1	598	10	O82747	O82747 arabidopsis
16	60	18.0	1802	5	Q9TXV8	Q9txv8 caenorhabdi
17	60	18.0	1880	13	Q9IBF1	Q9ibf1 takifugu pa
18	59.5	17.8	383	10	O82729	O82729 borago offi
19	59.5	17.8	473	5	Q94789	Q94789 trichostron
20	59.5	17.8	512	5	P92039	P92039 haemonchus
21	58.5	17.5	426	5	O96760	O96760 ascaris suu
22	58.5	17.5	586	10	Q9SNT3	Q9snt3 oryza sativ
23	58	17.4	250	2	Q9RSP1	Q9rsp1 deinococcus
24	57	17.1	1717	13	Q90519	Q90519 fugu rubrip
25	56.5	16.9	272	4	Q9UKB8	Q9ukb8 homo sapien
26	56.5	16.9	383	10	Q9LLL7	Q9lll7 sesamum ind
27	56.5	16.9	870	4	O60309	O60309 homo sapien
28	56.5	16.9	1400	5	Q20766	Q20766 caenorhabdi
29	56.5	16.9	1977	4	Q15858	Q15858 homo sapien
30	56.5	16.9	3268	3	Q03280	Q03280 saccharomyc
31	56	16.8	373	4	Q9NS66	Q9ns66 homo sapien
32	56	16.8	373	11	Q9JJH2	Q9jjh2 rattus norv
33	56	16.8	403	10	Q9LPG9	Q9lpg9 arabidopsis
34	56	16.8	512	2	Q9JVA8	Q9jva8 neisseria m
35	56	16.8	524	2	Q9K0A2	Q9k0a2 neisseria m
36	56	16.8	830	10	O65482	O65482 arabidopsis
37	56	16.8	1215	2	Q9Z771	Q9z771 chlamydia p
38	56	16.8	1215	2	Q9JS99	Q9js99 chlamydia p
39	55.5	16.6	372	5	Q18933	Q18933 caenorhabdi
40	55.5	16.6	378	10	O24499	O24499 helianthus
41	55.5	16.6	495	2	O53300	O53300 mycobacteri
42	55.5	16.6	706	5	Q9XUG5	Q9xug5 caenorhabdi
43	55.5	16.6	837	4	O75173	O75173 homo sapien
44	55.5	16.6	837	4	Q9UN83	Q9un83 homo sapien
45	55	16.5	276	11	Q60828	Q60828 mus musculu

ALIGNMENTS

RESULT 1
Q9NS76
ID Q9NS76 PRELIMINARY; PRT; 70 AA.
AC Q9NS76;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ABC1 (FRAGMENT).
GN ABC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Zhao L., Zhou C., Tanaka A., Nakata M., Hirabayashi T., Amachi T.,
RA Shioda S., Ueda K., Inagaki N.;
RT "Cloning, characterization, and tissue distribution of the ABC
RT transporter ABC2.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB037924; BAB07875.1; -.
FT NON_TER 70 70
SQ SEQUENCE 70 AA; 8383 MW; C6DBDEFE854F034F CRC64;

Query Match 100.0%; Score 334; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
```

RESULT 2
Q9NP93
ID Q9NP93 PRELIMINARY; PRT; 100 AA.
AC Q9NP93;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

```
DE   ATP BINDING CASSETTE TRANSPORTER 1 (FRAGMENT).
GN   ABCA1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Pullinger C.R., Hakamata H., Duchateau P.N., Eng C., Aouizerat B.E.,
RA   Fielding C.J., Kane J.P.;
RT   "Analysis of hABC1 gene 5' end: additional peptide sequence, promoter
RT   region, and four polymorphisms.";
RL   Biochem. Biophys. Res. Commun. 271:0-0(2000).
DR   EMBL; AF258626; AAF69516.1; -.
DR   EMBL; AF258624; AAF69516.1; JOINED.
DR   EMBL; AF258625; AAF69516.1; JOINED.
DR   EMBL; AF258627; AAF69513.1; -.
FT   NON_TER     100    100
SQ   SEQUENCE   100 AA;  11530 MW;  ABEBA02D542CE853 CRC64;


  Query Match             100.0%;  Score 334;  DB 4;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-34;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   3
Q9NQV4
ID   Q9NQV4      PRELIMINARY;      PRT;  2261 AA.
AC   Q9NQV4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ABCA1.
GN   ABCA1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Santamarina-Fojo S., Peterson K.M., Knapper C.L., Qiu Y.,
RA   Freeman L.A., Cheng J.-F., Osorio J., Remaley A.T., Yang X.-P.,
RA   Haudenschild C.C., Prades C., Chimini G., Blackmon E.E.,
RA   Francois T.L., Duverger N., Rubin E.M., Rosier M., Denefle P.,
RA   Fredrickson D.S., Brewer H.B. Jr.;
RT   "Complete genomic sequence of the human ABCA1 gene: Analysis of the
RT   human and mouse ATP-binding cassette A promoter.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:7987-7992(2000).
DR   EMBL; AF275948; AAF86276.1; -.
SQ   SEQUENCE   2261 AA;  254324 MW;  BA27D9B217ACAA33 CRC64;


  Query Match             100.0%;  Score 334;  DB 4;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-33;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   4
Q9NR73
ID   Q9NR73      PRELIMINARY;      PRT;  2146 AA.
AC   Q9NR73;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   MACROPHAGE ABC TRANSPORTER.
GN   ABCA7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kaminski W.E., Orso E., Diederich W., Klucken J., Drobnik W.,
RA   Schmitz G.;
RT   "Identification of a Novel Human Sterol-Sensitive ATP-Binding Cassette
RT   Transporter (ABCA7).";
RL   Biochem. Biophys. Res. Commun. 273:532-538(2000).
DR   EMBL; AF250238; AAF85794.1; -.
SQ   SEQUENCE   2146 AA;  234468 MW;  679B16EB2D75FF0D CRC64;


  Query Match             68.9%;  Score 230;  DB 4;  Length 2146;
  Best Local Similarity   67.8%;  Pred. No. 2.3e-20;
  Matches   40;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
          || | || |||||| :|||| |||:|: |||:|: ||::|| |:|| | ||||||||
Db      1 MAFWTQLMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 59


RESULT   5
O35600
ID   O35600      PRELIMINARY;      PRT;  2310 AA.
AC   O35600;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ATP-BINDING CASSETTE TRANSPORTER.
GN   ABCA4 OR ABC10 OR ABCR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=97345663; PubMed=9202155;
RA   Azarian S.M., Travis G.H.;
RT   "The photoreceptor rim protein is an ABC transporter encoded by the
RT   gene for recessive Stargardt's disease (ABCR).";
RL   FEBS Lett. 409:247-252(1997).
DR   EMBL; AF000149; AAC23916.1; -.
DR   MGD; MGI:109424; Abca4.
DR   INTERPRO; IPR001617; -.
DR   PFAM; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW   ATP-binding.
SQ   SEQUENCE   2310 AA;  260207 MW;  8370C6C8A62EF294 CRC64;


  Query Match             55.4%;  Score 185;  DB 11;  Length 2310;
  Best Local Similarity   58.2%;  Pred. No. 1e-14;
  Matches   32;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      6 QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          |::|||||| | |:|| : ::|: ||| :||:|| :| : |: | ||||||||||
Db      6 QIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYSQHECHFPNKA 60


RESULT   6
O02698
ID   O02698      PRELIMINARY;      PRT;  2281 AA.
AC   O02698;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
```

```
DE   ABC TRANSPORTER.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=RETINAL ROD;
RX   MEDLINE=97248596; PubMed=9092582;
RA   Illing M., Molday L.L., Molday R.S.;
RT   "The 220-kDa rim protein of retinal rod outer segments is a member of
RT   the ABC transporter superfamily.";
RL   J. Biol. Chem. 272:10303-10310(1997).
DR   EMBL; U90126; AAC48716.1; -.
DR   INTERPRO; IPR001617; -.
DR   PFAM; PF00005; ABC_tran; 2.
SQ   SEQUENCE   2281 AA;  257228 MW;  71CD404C98F7A079 CRC64;


  Query Match             53.6%;  Score 179;  DB 6;  Length 2281;
  Best Local Similarity   56.4%;  Pred. No. 5.8e-14;
  Matches   31;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       6 QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
           |::|||||| | |:||  : ::|: ||| :||:|| :|   | | :|||||||||
Db       6 QIKLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNVNPLYSKHECHFPNKA 60


RESULT   7
Q9N403
ID   Q9N403      PRELIMINARY;      PRT;   602 AA.
AC   Q9N403;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL PROTEIN Y39G10AL.B.
GN   Y39G10AL.B.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024766; AAF59489.1; -.
SQ   SEQUENCE   602 AA;  68965 MW;  B7E2A21F7D15C200 CRC64;


  Query Match             29.9%;  Score 100;  DB 5;  Length 602;
  Best Local Similarity   37.5%;  Pred. No. 0.00013;
  Matches   21;  Conservative    9;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       4 WPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
           | ||  |||| :  ||||   | :|:  |  :|:|:  ||         :||: :|
Db      42 WRQLNFLLWKCVLVRRRQKFWLAVELIVPCILFIIIALVRTKDFSEASPQCHYDSK 97


RESULT   8
Q9VRG4
ID   Q9VRG4      PRELIMINARY;      PRT;  1713 AA.
AC   Q9VRG4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG1718 PROTEIN.
GN   CG1718.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003568; AAF50837.1; -.
DR   FLYBASE; FBgn0031170; CG1718.
DR   INTERPRO; IPR001617; -.
DR   PFAM; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
SQ   SEQUENCE   1713 AA;  192888 MW;  9DE20D3BFB9DC1CA CRC64;


  Query Match             23.4%;  Score 78;  DB 5;  Length 1713;
  Best Local Similarity   37.5%;  Pred. No. 0.19;
  Matches   15;  Conservative    8;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVR 43
           | :  |||||| | :     |:::|:  |    |:|: ||
Db       7 WDKFVLLLWKNWTLQWNHKWQMVIELVLPAIFSLLLVLVR 46


RESULT   9
Q9T050
ID   Q9T050      PRELIMINARY;      PRT;   707 AA.
AC   Q9T050;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 79.4 KDA PROTEIN.
GN   T26M18.20 OR AT4G11810.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA   Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Mannhaupt G.,
RA   Schueller C.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA   Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL078606; CAB44319.1; -.
DR   EMBL; AL161532; CAB78224.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   707 AA;  79418 MW;  DA1F8F4BF88A7FF2 CRC64;

  Query Match             21.4%;  Score 71.5;  DB 10;  Length 707;
  Best Local Similarity   36.7%;  Pred. No. 0.56;
  Matches   22;  Conservative    7;  Mismatches   24;  Indels    7;  Gaps    2;

Qy       1 MACWPQLRLLLW-----KNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECH 55
           ||| | |  ||      ||:|| :     :: ||| |::  : || |   |  |  | |
Db     390 MACGPALAGLLQTDFKIKNVTFNQDTLPGWVMAVAWLLYLVWLAISFR--EPAREPEEIH 447


RESULT  10
Q9PVW2
ID   Q9PVW2      PRELIMINARY;      PRT;   307 AA.
AC   Q9PVW2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   OLFACTORY RECEPTOR 1.
GN   MFOR1.
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGY;
RA   Sun H., Kondo R., Shima A., Naruse K., Hori H., Chigusa S.I.;
RT   "Evolutionary analysis of putative olfactory receptor genes of medaka
RT   fish, Oryzias latipes.";
RL   Gene 231:137-145(1999).
DR   EMBL; AB022646; BAA84275.1; -.
DR   INTERPRO; IPR000276; -.
DR   PFAM; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEPTOR; UNKNOWN_1.
KW   Receptor.
SQ   SEQUENCE   307 AA;  35400 MW;  E14065B68FCCBEF8 CRC64;

  Query Match             19.6%;  Score 65.5;  DB 13;  Length 307;
  Best Local Similarity   33.3%;  Pred. No. 1.5;
  Matches   18;  Conservative    8;  Mismatches   21;  Indels    7;  Gaps    2;

Qy       3 CWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHF 56
           | |    | :::|  :|| |  |:|   ||||   : || ::      |  | |
Db     125 CKP----LQYQSLMSKRRVTVMLILSWLWPLFQLTVAISGKVI---INQRPCSF 171


RESULT  11
Q9PVW1
ID   Q9PVW1      PRELIMINARY;      PRT;   307 AA.
AC   Q9PVW1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   OLFACTORY RECEPTOR 2.
GN   MFOR2.
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGY;
RA   Sun H., Kondo R., Shima A., Naruse K., Hori H., Chigusa S.I.;
RT   "Evolutionary analysis of putative olfactory receptor genes of medaka
RT   fish, Oryzias latipes.";
RL   Gene 231:137-145(1999).
DR   EMBL; AB022647; BAA84276.1; -.
DR   INTERPRO; IPR000276; -.
DR   PFAM; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEPTOR; UNKNOWN_1.
KW   Receptor.
SQ   SEQUENCE   307 AA;  35678 MW;  BCA9C5FDF050B7D8 CRC64;

  Query Match             19.6%;  Score 65.5;  DB 13;  Length 307;
  Best Local Similarity   33.3%;  Pred. No. 1.5;
  Matches   18;  Conservative    8;  Mismatches   21;  Indels    7;  Gaps    2;

Qy       3 CWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHF 56
           | |    | :::|  :|| |  |:|   ||||   : || ::      |  | |
Db     125 CKP----LQYQSLMSKRRVTVMLILSWLWPLFQLTVAISGKVI---INQRPCSF 171


RESULT  12
Q9ZPB9
ID   Q9ZPB9      PRELIMINARY;      PRT;   129 AA.
AC   Q9ZPB9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   MATURASE (FRAGMENT).
GN   YCF14 OR MATK.
OS   Crawfurdia speciosa.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Gentianales; Gentianaceae; Crawfurdia.
OX   NCBI_TaxID=82711;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Thiv M., Kadereit J.W.;
RT   "The phylogenetic relationships and evolution of the Canarian laurel
RT   forest endemic Ixanthus viscosus (Ait.) Griseb. (Gentianaceae):
RT   evidence from matK and ITS sequences.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ010512; CAB37008.1; -.
DR   MENDEL; 40044; Crasp;ycf14;40044.
```

```
DR   INTERPRO; IPR002866; -.
DR   PFAM; PF01824; MatK_N; 1.
FT   NON_TER       1      1
FT   NON_TER     129    129
SQ   SEQUENCE   129 AA;  16059 MW;  4E86E1A8B9EC30DD CRC64;


  Query Match             18.4%;  Score 61.5;  DB 10;  Length 129;
  Best Local Similarity   25.0%;  Pred. No. 2.2;
  Matches   17;  Conservative    8;  Mismatches   20;  Indels   23;  Gaps    3;

Qy       1 MACW----PQLRLLL--------WKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPP 48
           : ||    | | ||         | ||: |:     || ::    |:||
Db      12 LRCWVKDTPSLHLLRFFFHEYWDWNNLSVRKNPGFALLTKINQRFFVFL---------- 60

Qy      49 YEQHECHF 56
           |  : | :
Db      61 YNSYVCEY 68


RESULT  13
Q9VZS7
ID   Q9VZS7      PRELIMINARY;      PRT;  2272 AA.
AC   Q9VZS7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG14967 PROTEIN.
GN   CG14967.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003477; AAF47740.1; -.
DR   FLYBASE; FBgn0035420; CG14967.
SQ   SEQUENCE   2272 AA;  259619 MW;  4102D7EC483A3298 CRC64;


  Query Match             18.4%;  Score 61.5;  DB 5;  Length 2272;
  Best Local Similarity   27.5%;  Pred. No. 28;
  Matches   22;  Conservative    9;  Mismatches   26;  Indels   23;  Gaps    2;

Qy       2 ACWPQLRLLLWKNLTFRRRQTCQLL-----------LEVAWP----------LFIFLI 38
           |||  |||||   ||   :|   ||            :|: |           :|   :
Db    1142 ACWEPLRLLLHGRLTLIAKQFTILLHASLDPYNTTEEMELTWNNCGIVLTNAKIMFKGEL 1201

Qy      39 LISVRLSYPPYEQHECHFPN 58
             ::|| :    :    ||||
Db    1202 NVTVRTASRYDDCRLLHFPN 1221


RESULT  14
Q9NF52
ID   Q9NF52      PRELIMINARY;      PRT;   218 AA.
AC   Q9NF52;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   BACN5I9.GS.1.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Murphy L., Harris D., Barrell B.;
RT   "Sequencing the distal X chromosome of Drosophila melanogaster.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Benos P.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL121800; CAB58006.1; -.
SQ   SEQUENCE   218 AA;  25502 MW;  E9DF6B80F0356CBA CRC64;


  Query Match             18.3%;  Score 61;  DB 5;  Length 218;
  Best Local Similarity   23.3%;  Pred. No. 4;
  Matches   14;  Conservative   17;  Mismatches   27;  Indels    2;  Gaps    2;

Qy       1 MACWPQLRLLLWKNLTFRRRQTCQL-LLEVAWPLFIF-LILISVRLSYPPYEQHECHFPN 58
           :| | |    : |: | |: :   | :: :|| : |: :: : : | :   :  : | ||
Db      90 LAMWRQRHFYVLKHWTLRKARLLSLAIILIAWLMLIYGIVTVDLPLRHVAMDSGDQHRPN 149


RESULT  15
O82747
ID   O82747      PRELIMINARY;      PRT;   598 AA.
AC   O82747;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 67.7 KDA PROTEIN.
GN   F7H19.170 OR AT4G22990.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Bancroft I.,
```

```
RA   Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL031018; CAA19814.1; -.
DR   EMBL; AL161558; CAB79254.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   598 AA;  67750 MW;  89F8821BC5081826 CRC64;


  Query Match             18.1%;  Score 60.5;  DB 10;  Length 598;
  Best Local Similarity   30.5%;  Pred. No. 11;
  Matches   18;  Conservative   11;  Mismatches   21;  Indels    9;  Gaps    3;

Qy       1 MACWP------QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHE 53
           ||| |      |:|  ::| ||| :      :: |||   |:|: :::    |  |  |
Db     387 MACGPALAGLLQIRFKIYK-LTFNQDTLPGWVMAVAW--LIYLVWLAISFREPAREPEE 442


Search completed: May 31, 2001, 13:09:59
Job time: 325 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 31, 2001, 13:03:59 ; Search time 31.1 Seconds
(without alignments)
66.088 Million cell updates/sec

Title: US-09-526-193A-1_COPY_1_60
Perfect score: 334
Sequence: 1 MACWPQLRLLLWKNLTFRRR.....SVRLSYPPYEQHECHFPNKA 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	334	100.0	2261	1	ABC1_HUMAN	O95477 homo sapien
2	334	100.0	2261	1	ABC1_MOUSE	P41233 mus musculu
3	180	53.9	2273	1	ABCR_HUMAN	P78363 homo sapien
4	84	25.1	1704	1	ABC3_HUMAN	Q99758 homo sapien
5	65.5	19.6	778	1	FTSK_COXBU	P39920 coxiella bu
6	60.5	18.1	511	1	ACH5_CAEEL	Q23022 caenorhabdi
7	58.5	17.5	383	1	FD62_SOYBN	P48631 glycine max
8	57	17.1	1093	1	SYV_NEUCR	P28350 neurospora
9	57	17.1	1103	1	CYGF_BOVIN	O02740 bos taurus
10	56.5	16.9	870	1	Y563_HUMAN	O60309 homo sapien
11	56	16.8	463	1	D5DR_FUGRU	P53454 fugu rubrip
12	56	16.8	921	1	CN2A_BOVIN	P14099 bos taurus
13	56	16.8	941	1	CN2A_HUMAN	O00408 homo sapien
14	55.5	16.6	436	1	ACHX_ONCVO	P54247 onchocerca
15	55.5	16.6	837	1	ATS4_HUMAN	O75173 homo sapien
16	55	16.5	357	1	YCFT_ECOLI	P75955 escherichia
17	55	16.5	383	1	FD6E_ARATH	P46313 arabidopsis
18	55	16.5	928	1	CN2A_RAT	Q01062 rattus norv
19	54	16.2	2005	1	CIN2_HUMAN	Q99250 homo sapien
20	54	16.2	2005	1	CIN2_RAT	P04775 rattus norv
21	53.5	16.0	333	1	CXC1_HUMAN	P46094 homo sapien
22	53.5	16.0	1682	1	CIN6_HUMAN	Q01118 homo sapien
23	53	15.9	475	1	DXR_MENPI	Q9xes0 mentha pipe
24	53	15.9	638	1	NTGL_BOVIN	Q28039 bos taurus
25	53	15.9	1229	1	CY43_TRYBB	Q99280 trypanosoma
26	52.5	15.7	282	1	NU2M_CAEEL	P24889 caenorhabdi
27	52.5	15.7	345	1	C5AR_CAVPO	O70129 cavia porce
28	52.5	15.7	434	1	FES_ERWCH	O51900 erwinia chr
29	52.5	15.7	465	1	DCDR_XENLA	P42291 xenopus lae
30	52.5	15.7	676	1	EXL1_HUMAN	Q92935 homo sapien
31	52	15.6	378	1	EXT1_CAEEL	O01704 caenorhabdi
32	52	15.6	633	1	NTGL_MOUSE	P28571 mus musculu
33	52	15.6	633	1	NTGL_RAT	P28572 rattus norv
34	52	15.6	692	1	NTGL_HUMAN	P48067 homo sapien
35	52	15.6	1951	1	CIN3_HUMAN	Q9ny46 homo sapien
36	52	15.6	1951	1	CIN3_RAT	P08104 rattus norv
37	52	15.6	2016	1	CIN5_HUMAN	Q14524 homo sapien
38	52	15.6	2019	1	CIN5_RAT	P15389 rattus norv
39	51.5	15.4	244	1	YRT2_CAEEL	Q10045 caenorhabdi
40	51.5	15.4	322	1	CXC1_MOUSE	Q9r0m1 mus musculu
41	51.5	15.4	378	1	EDG3_HUMAN	Q99500 homo sapien
42	51.5	15.4	989	1	PRTH_PORGI	P46071 porphyromon
43	51.5	15.4	2009	1	CIN1_HUMAN	P35498 homo sapien
44	51.5	15.4	2009	1	CIN1_RAT	P04774 rattus norv
45	51	15.3	293	1	URER_PROMI	Q02458 proteus mir

ALIGNMENTS

RESULT 1
ABC1_HUMAN
ID ABC1_HUMAN STANDARD; PRT; 2261 AA.
AC O95477; Q9UN08; Q9UN07; Q9UN06; Q9NQV4; Q9UN09;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ATP-BINDING CASSETTE, SUB-FAMILY A, MEMBER 1 (ATP-BINDING CASSETTE
DE TRANSPORTER 1) (ATP-BINDING CASSETTE 1) (ABC-1) (CHOLESTEROL EFFLUX
DE REGULATORY PROTEIN).
GN ABCA1 OR ABC1 OR CERP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20345099; PubMed=10884428;
RA Santamarina-Fojo S., Peterson K.M., Knapper C.L., Qiu Y.,
RA Freeman L.A., Cheng J.-F., Osorio J., Remaley A.T., Yang X.-P.,
RA Haudenschild C.C., Prades C., Chimini G., Blackmon E.E.,
RA Francois T.L., Duverger N., Rubin E.M., Rosier M., Denefle P.,
RA Fredrickson D.S., Brewer H.B. Jr.;
RT "Complete genomic sequence of the human ABCA1 gene: analysis of the
RT human and mouse ATP-binding cassette A promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7987-7992(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Schwartz K., Lawn R.M., Wade D.P.;
RT "ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
RT regulated by LXR.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 21-2261 FROM N.A.
RX PubMed=10092505;
RA Langmann T., Klucken J., Reil M., Liebisch G., Luciani M.F.,
RA Chimini G., Kaminski W., Schmitz G.;
RT "Molecular cloning of the human ATP-binding cassette transporter 1
RT (hABC1): evidence for dterol-dependent regulation in macrophages.";
RL Biochem. Biophys. Res. Commun. 257:29-33(1999).
RN [4]
RP SEQUENCE OF 21-2261 FROM N.A.
RX MEDLINE=99364413; PubMed=10431238;
RA Rust S., Rosier M., Funke H., Real J., Amoura Z., Piette J.-C.,
RA Deleuze J.-F., Brewer H.B., Duverger N., Denefle P., Assmann G.;
RT "Tangier disease is caused by mutations in the gene encoding
RT ATP-binding cassette transporter 1.";
RL Nat. Genet. 22:352-355(1999).
RN [5]
RP VARIANTS TD ARG-597; LEU-693 DEL AND ARG-1477.
RX MEDLINE=99364411; PubMed=10431236;
RA Brooks-Wilson A., Marcil M., Clee S.M., Zhang L.-H., Roomp K.,
RA van Dam M., Yu L., Brewer C., Collins J.A., Molhuizen H.O.F.,
RA Loubser O., Ouelette B.F.F., Fichter K., Ashbourne-Excoffon K.J.,
RA Sensen C.W., Scherer S., Mott S., Denis M., Martindale D.,

```
RA   Frohlich J., Morgan K., Koop B., Pimstone S., Kastelein J.J.,
RA   Hayden M.R.;
RT   "Mutations in ABC1 in Tangier disease and familial high-density
RT   lipoprotein deficiency.";
RL   Nat. Genet. 22:336-345(1999).
RN   [6]
RP   VARIANTS TD SER-590; SER-935 AND VAL-937.
RX   MEDLINE=99364412; PubMed=10431237;
RA   Bodzioch M., Orso E., Klucken J., Langmann T., Bottcher A.,
RA   Diederich W., Drobnik W., Barlage S., Buchler C., Porsch-Ozcurumez M.,
RA   Kaminski W.E., Hahmann H.W., Oette K., Rothe G., Aslanidis C.,
RA   Lackner K.J., Schmitz G.;
RT   "The gene encoding ATP-binding cassette transporter 1 is mutated in
RT   Tangier disease.";
RL   Nat. Genet. 22:347-351(1999).
RN   [7]
RP   VARIANTS TD LEU-1289 AND HIS-1800.
RX   MEDLINE=20171564; PubMed=10706591;
RA   Brousseau M.E., Schaefer E.J., Dupuis J., Eustace B.,
RA   Van Eerdewegh P., Goldkamp A.L., Thurston L.M., FitzGerald M.G.,
RA   Yasek-McKenna D., O'Neill G., Eberhart G.P., Weiffenbach B.,
RA   Ordovas J.M., Freeman M.W., Brown R.H. Jr., Gu J.Z.;
RT   "Novel mutations in the gene encoding ATP-binding cassette 1 in four
RT   tangier disease kindreds.";
RL   J. Lipid Res. 41:433-441(2000).
CC   -!- FUNCTION: CAMP-DEPENDENT AND SULFONYLUREA-SENSITIVE ANION
CC       TRANSPORTER. KEY GATEKEEPER INFLUENCING INTRACELLULAR CHOLESTEROL
CC       TRANSPORT.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED, BUT MOST ABUNDANT IN
CC       MACROPHAGES.
CC   -!- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC       EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC       ATP BINDING CASSETTE (ABC) DOMAIN.
CC   -!- DISEASE: DEFECTS IN ABCA1 ARE A CAUSE OF TANGIER DISEASE (TD). TD
CC       IS A RECESSIVE DISORDER CHARACTERIZED BY ABSENCE OF HIGH DENSITY
CC       LIPOPROTEIN (HDL) CHOLESTEROL FROM PLASMA, HEPATOSPLENOMEGALY,
CC       PERIPHERAL NEUROPATHY, AND FREQUENTLY PREMATURE CORONARY ARTERY
CC       DISEASE (CHD).
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS). MDR SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF275948; AAF86276.1; -.
DR   EMBL; AF285167; AAF98175.1; -.
DR   EMBL; AJ012376; CAA10005.1; ALT_INIT.
DR   EMBL; AF165281; AAD49849.1; ALT_INIT.
DR   EMBL; AF165286; AAD49851.1; -.
DR   EMBL; AF165282; AAD49851.1; JOINED.
DR   EMBL; AF165283; AAD49851.1; JOINED.
DR   EMBL; AF165284; AAD49851.1; JOINED.
DR   EMBL; AF165285; AAD49851.1; JOINED.
DR   EMBL; AF165306; AAD49852.1; -.
DR   EMBL; AF165287; AAD49852.1; JOINED.
DR   EMBL; AF165288; AAD49852.1; JOINED.
DR   EMBL; AF165289; AAD49852.1; JOINED.
DR   EMBL; AF165290; AAD49852.1; JOINED.
DR   EMBL; AF165291; AAD49852.1; JOINED.
DR   EMBL; AF165292; AAD49852.1; JOINED.
DR   EMBL; AF165293; AAD49852.1; JOINED.
DR   EMBL; AF165294; AAD49852.1; JOINED.
DR   EMBL; AF165295; AAD49852.1; JOINED.
DR   EMBL; AF165296; AAD49852.1; JOINED.
DR   EMBL; AF165297; AAD49852.1; JOINED.
DR   EMBL; AF165298; AAD49852.1; JOINED.
DR   EMBL; AF165299; AAD49852.1; JOINED.
DR   EMBL; AF165300; AAD49852.1; JOINED.
DR   EMBL; AF165301; AAD49852.1; JOINED.
DR   EMBL; AF165302; AAD49852.1; JOINED.
DR   EMBL; AF165303; AAD49852.1; JOINED.
DR   EMBL; AF165304; AAD49852.1; JOINED.
DR   EMBL; AF165305; AAD49852.1; JOINED.
DR   EMBL; AF165309; AAD49854.1; -.
DR   EMBL; AF165307; AAD49854.1; JOINED.
DR   EMBL; AF165308; AAD49854.1; JOINED.
DR   EMBL; AF165310; AAD49853.1; -.
DR   MIM; 600046; -.
DR   MIM; 205400; -.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Glycoprotein; Transmembrane; Transport;
KW   Disease mutation.
FT   TRANSMEM     26     42       POTENTIAL.
FT   TRANSMEM    640    656       POTENTIAL.
FT   TRANSMEM    690    706       POTENTIAL.
FT   TRANSMEM    717    733       POTENTIAL.
FT   TRANSMEM    749    765       POTENTIAL.
FT   TRANSMEM    771    787       POTENTIAL.
FT   TRANSMEM   1041   1057       POTENTIAL.
FT   TRANSMEM   1351   1367       POTENTIAL.
FT   TRANSMEM   1661   1677       POTENTIAL.
FT   TRANSMEM   1708   1724       POTENTIAL.
FT   TRANSMEM   1737   1753       POTENTIAL.
FT   TRANSMEM   1775   1791       POTENTIAL.
FT   TRANSMEM   1854   1870       POTENTIAL.
FT   NP_BIND     933    940       ATP (POTENTIAL).
FT   NP_BIND    1946   1953       ATP (POTENTIAL).
FT   CARBOHYD     14     14       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     98     98       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    196    196       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    337    337       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    478    478       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    489    489       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    521    521       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    820    820       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1144   1144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1294   1294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1453   1453       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1504   1504       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1637   1637       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2044   2044       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2238   2238       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     399    399       V -> A (IN TD).
FT                                /FTId=VAR_009145.
FT   VARIANT     587    587       R -> W (IN TD).
FT                                /FTId=VAR_009146.
FT   VARIANT     590    590       W -> S (IN TD).
FT                                /FTId=VAR_009147.
FT   VARIANT     597    597       Q -> R (IN TD).
FT                                /FTId=VAR_009148.
FT   VARIANT     693    693       MISSING (IN TD).
FT                                /FTId=VAR_009149.
FT   VARIANT     935    935       N -> S (IN TD).
FT                                /FTId=VAR_009150.
FT   VARIANT     937    937       A -> V (IN TD).
FT                                /FTId=VAR_009151.
FT   VARIANT    1289   1289       D -> L (IN TD).
FT                                /FTId=VAR_009152.
FT   VARIANT    1477   1477       C -> R (IN TD).
FT                                /FTId=VAR_009153.
FT   VARIANT    1517   1517       I -> R (IN TD).
FT                                /FTId=VAR_009154.
FT   VARIANT    1800   1800       N -> H (IN TD).
```

```
  Query Match             100.0%;  Score 334;  DB 1;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-31;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   2
ABC1_MOUSE
ID   ABC1_MOUSE     STANDARD;      PRT;  2261 AA.
AC   P41233;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ATP-BINDING CASSETTE, SUB-FAMILY A, MEMBER 1 (ATP-BINDING CASSETTE
DE   TRANSPORTER 1) (ATP-BINDING CASSETTE 1) (ABC-1).
GN   ABCA1 OR ABC1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2; TISSUE=Macrophage;
RX   MEDLINE=94375008; PubMed=8088782;
RA   Luciani M.F., Denizot F., Savary S., Mattei M.-G., Chimini G.;
RT   "Cloning of two novel ABC transporters mapping on human chromosome
RT   9.";
RL   Genomics 21:150-159(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Qiu Y., Cavelier L., Chiu S., Rubin E., Cheng J.-F.;
RT   "Human and mouse ABCA1 comparative sequencing and transgenesis studies
RT   identify potential regulatory sequences.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CAMP-DEPENDENT AND SULFONYLUREA-SENSITIVE ANION
CC       TRANSPORTER. KEY GATEKEEPER INFLUENCING INTRACELLULAR CHOLESTEROL
CC       TRANSPORT (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN ADULT TISSUES. HIGHEST
CC       LEVELS ARE FOUND IN PREGNANT UTERUS AND UTERUS.
CC   -!- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC       EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC       ATP BINDING CASSETTE (ABC) DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS). MDR SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X75926; CAA53530.1; ALT_INIT.
DR   EMBL; AF287263; AAG39073.1; ALT_INIT.
DR   MGD; MGI:99607; Abca1.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Glycoprotein; Transmembrane; Transport.
FT   TRANSMEM     26     42       POTENTIAL.
FT   TRANSMEM    640    656       POTENTIAL.
FT   TRANSMEM    690    706       POTENTIAL.
FT   TRANSMEM    717    733       POTENTIAL.
FT   TRANSMEM    749    765       POTENTIAL.
FT   TRANSMEM    771    787       POTENTIAL.
FT   TRANSMEM   1041   1057       POTENTIAL.
FT   TRANSMEM   1351   1367       POTENTIAL.
FT   TRANSMEM   1661   1677       POTENTIAL.
FT   TRANSMEM   1708   1724       POTENTIAL.
FT   TRANSMEM   1737   1753       POTENTIAL.
FT   TRANSMEM   1775   1791       POTENTIAL.
FT   TRANSMEM   1854   1870       POTENTIAL.
FT   NP_BIND     933    940       ATP (POTENTIAL).
FT   NP_BIND    1946   1953       ATP (POTENTIAL).
FT   CARBOHYD     14     14       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     98     98       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    196    196       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    244    244       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    337    337       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    478    478       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    489    489       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    521    521       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    820    820       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1144   1144       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1294   1294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1453   1453       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1499   1499       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1504   1504       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1637   1637       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2044   2044       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2238   2238       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   1567   1568       MISSING (IN REF. 2).
FT   CONFLICT   2024   2024       MISSING (IN REF. 2).
SQ   SEQUENCE   2261 AA;  254011 MW;  FAE62B21FD1D09F9 CRC64;


  Query Match             100.0%;  Score 334;  DB 1;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-31;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   3
ABCR_HUMAN
ID   ABCR_HUMAN     STANDARD;      PRT;  2273 AA.
AC   P78363; O60438; O60915; O15112;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RETINAL-SPECIFIC ATP-BINDING CASSETTE TRANSPORTER (RIM ABC
DE   TRANSPORTER) (RIM PROTEIN) (RMP) (STARGARDT DISEASE PROTEIN).
GN   ABCA4 OR ABCR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS STGD.
RX   MEDLINE=97207641; PubMed=9054934;
RA   Allikmets R., Singh N., Sun H., Shroyer N.F., Hutchinson A.,
RA   Chidambaram A., Gerrard B., Baird L., Stauffer D., Peiffer A.,
RA   Rattner A., Smallwood P., Li Y., Anderson K.L., Lewis R.A.,
RA   Nathans J., Leppert M., Dean M., Lupski J.R.;
RT   "A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is
RT   mutated in recessive Stargardt macular dystrophy.";
RL   Nat. Genet. 15:236-246(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97345663; PubMed=9202155;
RA   Azarian S.M., Travis G.H.;
RT   "The photoreceptor rim protein is an ABC transporter encoded by the
RT   gene for recessive Stargardt's disease (ABCR).";
```

RL FEBS Lett. 409:247-252(1997).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS STGD W-18 AND C-212.
RX MEDLINE=98163759; PubMed=9503029;
RA Gerber S., Rozet J.M., van de Pol T.J.R., Hoying C.B., Munnich A.,
RA Blankenagel A., Kaplan J., Cremers F.P.M.;
RT "Complete exon-intron structure of the retina-specific ATP binding
RT transporter gene (ABCR) allows the identification of novel mutations
RT underlying Stargardt disease.";
RL Genomics 48:139-142(1998).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANTS STGD.
RX MEDLINE=98141123; PubMed=9490294;
RA Nasonkin I., Illing M., Koehler M.R., Schmid M., Molday R.S.,
RA Weber B.H.;
RT "Mapping of the rod photoreceptor ABC transporter (ABCR) to 1p21-p22.1
RT and identification of novel mutations in Stargardt's disease.";
RL Hum. Genet. 102:21-26(1998).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=99175213; PubMed=10075733;
RA Sun H., Molday R.S., Nathans J.;
RT "Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR,
RT the photoreceptor-specific ATP-binding cassette transporter
RT responsible for Stargardt disease.";
RL J. Biol. Chem. 274:8269-8281(1999).
RN [6]
RP VARIANTS ARMD2.
RX MEDLINE=97442530; PubMed=9295268;
RA Allikmets R., Shroyer N.F., Singh N., Seddon J.M., Lewis R.A.,
RA Bernstein P.S., Peiffer A., Zabriskie N.A., Li Y., Hutchinson A.,
RA Dean M., Lupski J.R., Leppert M.;
RT "Mutation of the Stargardt disease gene (ABCR) in age-related macular
RT degeneration.";
RL Science 277:1805-1807(1997).
RN [7]
RP VARIANTS STGD.
RX MEDLINE=98454319; PubMed=9781034;
RA Rozet J.M., Gerber S., Souied E., Perrault I., Chatelin S., Ghazi I.,
RA Leowski C., Dufier J.L., Munnich A., Kaplan J.;
RT "Spectrum of ABCR gene mutations in autosomal recessive macular
RT dystrophies.";
RL Eur. J. Hum. Genet. 6:291-295(1998).
RN [8]
RP VARIANTS STGD.
RX MEDLINE=99138655; PubMed=9973280;
RA Lewis R.A., Shroyer N.F., Singh N., Allikmets R., Hutchinson A.,
RA Li Y., Lupski J.R., Leppert M., Dean M.;
RT "Genotype/phenotype analysis of a photoreceptor-specific ATP-binding
RT cassette transporter gene, ABCR, in Stargardt disease.";
RL Am. J. Hum. Genet. 64:422-434(1999).
RN [9]
RP VARIANTS STGD, AND VARIANTS.
RX MEDLINE=99192348; PubMed=10090887;
RA Maugeri A., van Driel M.A., van de Pol D.J.R., Klevering B.J.,
RA van Haren F.J.J., Tijmes N., Bergen A.A.B., Rohrschneider K.,
RA Blankenagel A., Pinckers A.J.L.G., Dahl N., Brunner H.G.,
RA Deutman A.F., Hoyng C.B., Cremers F.P.M.;
RT "The 2588G-->C mutation in the ABCR gene is a mild frequent founder
RT mutation in the western European population and allows the
RT classification of ABCR Mutations in patients with Stargardt disease.";
RL Am. J. Hum. Genet. 64:1024-1035(1999).
RN [10]
RP VARIANT STGD TYR-54, AND VARIANT ALA-863.
RX MEDLINE=20077755; PubMed=10612508;
RA Zhang K., Garibaldi D.C., Kniazeva M., Albini T., Chiang M.F.,
RA Kerrigan M., Sunness J.S., Han M., Allikmets R.;
RT "A novel mutation in the ABCR gene in four patients with autosomal
RT recessive Stargardt disease.";
RL Am. J. Ophthalmol. 128:720-724(1999).
RN [11]
RP VARIANTS STGD.
RX MEDLINE=20098082; PubMed=10634594;
RA Papaioannou M., Ocaka L., Bessant D., Lois N., Bird A., Payne A.,
RA Bhattacharya S.;
RT "An analysis of ABCR mutations in British patients with recessive
RT retinal dystrophies.";
RL Invest. Ophthalmol. Vis. Sci. 41:16-19(2000).
CC -!- FUNCTION: MAY PLAY A ROLE IN PHOTORESPONSE. RETINOIDS, AND MOST
CC LIKELY RETINAL, ARE THE NATURAL SUBSTRATES FOR TRANSPORT BY ABCR
CC IN ROD OUTER SEGMENTS. MAY ACT IN THE VISUAL CYCLE TO FLIP PE-ALL-
CC TRANS-RETINAL ADDUCTS FROM THE LUMENAL TO THE CYTOSOLIC FACE OF
CC THE DISK MEMBRANE, MOVE FREE ALL-TRANS-RETINAL FROM THE LIPID
CC PHASE OF THE DISC MEMBRANE TO A JUXTAMEMBRANE LOCATION, OR
CC POSSIBLY REORIENT ALL-TRANS-RETINAL IN THE BILAYER.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: RETINAL-SPECIFIC. SEEMS TO BE EXCLUSIVELY
CC FOUND IN THE RIMS OF ROD PHOTORECEPTOR CELLS.
CC -!- DISEASE: DEFECTS IN ABCA4 ARE A CAUSE OF STARGARDT DISEASE (STGD);
CC ALSO KNOWN AS FUNDUS FLAVIMACULATUS (FFM) OR JUVENILE MACULAR
CC DEGENERATION. IT IS AN AUTOSOMAL RECESSIVE RETINAL DISORDER. IT IS
CC ONE OF THE MOST FREQUENT CAUSES OF MACULAR DEGENERATION IN
CC CHILDHOOD. IT IS CHARACTERIZED BY A JUVENILE-ONSET MACULAR
CC DYSTROPHY, ALTERATIONS OF THE PERIPHERAL RETINA, AND SUBRETINAL
CC DEPOSITION OF LIPOFUSCIN-LIKE MATERIAL.
CC -!- DISEASE: DEFECTS IN ABCA4 ARE A CAUSE OF AGE-RELATED MACULAR
CC DEGENERATION (ARMD2 OR AMD).
CC -!- DISEASE: DEFECTS IN ABCA4 ARE A CAUSE OF AUTOSOMAL RECESSIVE CONE-
CC ROD DISTROPHY (ARCRD OR CRD).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
CC -!- DATABASE: NAME=Mutations of the ABCA4 gene;
CC NOTE=Retina International's Scientific Newsletter;
CC WWW="http://www.irpa.org/sci-news/abcrmut.htm".
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; U88667; AAC51144.1; -.
DR EMBL; AF000148; AAC23915.1; -.
DR EMBL; Y15635; CAA75729.1; -.
DR EMBL; Y15636; CAA75729.1; JOINED.
DR EMBL; Y15637; CAA75729.1; JOINED.
DR EMBL; Y15638; CAA75729.1; JOINED.
DR EMBL; Y15639; CAA75729.1; JOINED.
DR EMBL; Y15640; CAA75729.1; JOINED.
DR EMBL; Y15641; CAA75729.1; JOINED.
DR EMBL; Y15642; CAA75729.1; JOINED.
DR EMBL; Y15643; CAA75729.1; JOINED.
DR EMBL; Y15644; CAA75729.1; JOINED.
DR EMBL; Y15645; CAA75729.1; JOINED.
DR EMBL; Y15646; CAA75729.1; JOINED.
DR EMBL; Y15647; CAA75729.1; JOINED.
DR EMBL; Y15648; CAA75729.1; JOINED.
DR EMBL; Y15649; CAA75729.1; JOINED.
DR EMBL; Y15650; CAA75729.1; JOINED.
DR EMBL; Y15651; CAA75729.1; JOINED.
DR EMBL; Y15652; CAA75729.1; JOINED.
DR EMBL; Y15653; CAA75729.1; JOINED.
DR EMBL; Y15654; CAA75729.1; JOINED.
DR EMBL; Y15655; CAA75729.1; JOINED.
DR EMBL; Y15656; CAA75729.1; JOINED.
DR EMBL; Y15657; CAA75729.1; JOINED.
DR EMBL; Y15658; CAA75729.1; JOINED.
DR EMBL; Y15659; CAA75729.1; JOINED.
DR EMBL; Y15660; CAA75729.1; JOINED.
DR EMBL; Y15661; CAA75729.1; JOINED.
DR EMBL; Y15662; CAA75729.1; JOINED.
DR EMBL; Y15663; CAA75729.1; JOINED.
DR EMBL; Y15664; CAA75729.1; JOINED.
DR EMBL; Y15665; CAA75729.1; JOINED.

```
DR   EMBL; Y15666; CAA75729.1; JOINED.
DR   EMBL; Y15667; CAA75729.1; JOINED.
DR   EMBL; Y15668; CAA75729.1; JOINED.
DR   EMBL; Y15669; CAA75729.1; JOINED.
DR   EMBL; Y15670; CAA75729.1; JOINED.
DR   EMBL; Y15671; CAA75729.1; JOINED.
DR   EMBL; Y15672; CAA75729.1; JOINED.
DR   EMBL; Y15673; CAA75729.1; JOINED.
DR   EMBL; Y15674; CAA75729.1; JOINED.
DR   EMBL; Y15675; CAA75729.1; JOINED.
DR   EMBL; Y15676; CAA75729.1; JOINED.
DR   EMBL; Y15677; CAA75729.1; JOINED.
DR   EMBL; Y15678; CAA75729.1; JOINED.
DR   EMBL; Y15679; CAA75729.1; JOINED.
DR   EMBL; Y15680; CAA75729.1; JOINED.
DR   EMBL; Y15681; CAA75729.1; JOINED.
DR   EMBL; Y15682; CAA75729.1; JOINED.
DR   EMBL; Y15683; CAA75729.1; JOINED.
DR   EMBL; Y15684; CAA75729.1; JOINED.
DR   EMBL; AF001945; AAC05632.1; -.
DR   MIM; 601691; -.
DR   MIM; 248200; -.
DR   MIM; 153800; -.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.

  Query Match             53.9%;  Score 180;  DB 1;  Length 2273;
  Best Local Similarity   56.4%;  Pred. No. 1.8e-13;
  Matches   31;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      6 QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          |::|||||| | |:|| : ::|: ||| :||:|| :| : | |  |||||||||
Db      6 QIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYSHHECHFPNKA 60


RESULT   4
ABC3_HUMAN
ID   ABC3_HUMAN     STANDARD;      PRT;  1704 AA.
AC   Q99758; Q92473;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   ATP-BINDING CASSETTE, SUB-FAMILY A, MEMBER 3 (ATP-BINDING CASSETTE
DE   TRANSPORTER 3) (ATP-BINDING CASSETTE 3) (ABC-C TRANSPORTER).
GN   ABCA3 OR ABC3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thyroid carcinoma;
RX   MEDLINE=96326608; PubMed=8706931;
RA   Klugbauer N., Hofmann F.;
RT   "Primary structure of a novel ABC transporter with a chromosomal
RT   localization on the band encoding the multidrug resistance-associated
RT   protein.";
RL   FEBS Lett. 391:61-65(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97179225; PubMed=9027511;
RA   Connors T.D., van Raay T.J., Petry L.R., Klinger K.W., Landes G.M.,
RA   Burn T.C.;
RT   "The cloning of a human ABC gene (ABC3) mapping to chromosome
RT   16p13.3.";
RL   Genomics 39:231-234(1997).
CC   -!- FUNCTION: MAY BE A TRANSPORTER, ITS NATURAL SUBSTRATE HAS NOT BEEN
CC       FOUND YET (BY SIMILARITY). MAY ACT AS AN EFFLUX PUMP FOR
CC       CHEMOTHERAPEUTICS DRUGS.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, FOLLOWED BY BRAIN,
CC       PANCREAS, SKELETAL MUSCLE AND HEART. WEAKLY EXPRESSED IN PLACENTA,
CC       KIDNEY AND LIVER. ALSO EXPRESSED IN MEDULLARY THYROID CARCINOMA
CC       CELLS (MTC) AND IN C-CELL CARCINOMA.
CC   -!- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC       EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC       ATP BINDING CASSETTE (ABC) DOMAIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS). MDR SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U78735; AAC50967.1; -.
DR   EMBL; X97187; CAA65825.1; -.
DR   MIM; 601615; -.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transport; Transmembrane.
FT   TRANSMEM     22     42       POTENTIAL.
FT   TRANSMEM    249    269       POTENTIAL.
FT   TRANSMEM    307    327       POTENTIAL.
FT   TRANSMEM    344    364       POTENTIAL.
FT   TRANSMEM    373    393       POTENTIAL.
FT   TRANSMEM    405    425       POTENTIAL.
FT   TRANSMEM    447    467       POTENTIAL.
FT   TRANSMEM    925    945       POTENTIAL.
FT   TRANSMEM   1100   1120       POTENTIAL.
FT   TRANSMEM   1144   1164       POTENTIAL.
FT   TRANSMEM   1183   1203       POTENTIAL.
FT   TRANSMEM   1213   1233       POTENTIAL.
FT   TRANSMEM   1245   1265       POTENTIAL.
FT   TRANSMEM   1306   1326       POTENTIAL.
FT   NP_BIND     566    573       ATP (POTENTIAL).
FT   NP_BIND    1416   1423       ATP (POTENTIAL).
FT   CONFLICT     36     36       P -> S (IN REF. 2).
FT   CONFLICT    196    196       L -> P (IN REF. 2).
SQ   SEQUENCE   1704 AA;  191387 MW;  AF0098DAF7A04F5F CRC64;


  Query Match             25.1%;  Score 84;  DB 1;  Length 1704;
  Best Local Similarity   47.7%;  Pred. No. 0.026;
  Matches   21;  Conservative    6;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL 44
            ||   || |||||| | ::|:    :||:  ||    ||| :||
Db      1 MAVLRQLALLLWKNYTLQKRKVLVTVLELFLPLLFPGILIWLRL 44


RESULT   5
FTSK_COXBU
ID   FTSK_COXBU     STANDARD;      PRT;   778 AA.
AC   P39920;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CELL DIVISION PROTEIN FTSK HOMOLOG.
GN   FTSK OR SPOIIIE.
OS   Coxiella burnetii.
OC   Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC   Coxiella group; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NINE MILE PHASE I / BRATISLAVA;
RX   MEDLINE=94055499; PubMed=8237209;
RA   Oswald W., Thiele D.;
RT   "A sporulation gene in Coxiella burnetii?";
RL   J. Vet. Med. B 40:366-370(1993).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
```

```
CC   -!- SIMILARITY: BELONGS TO THE FTSK/SPOIIIE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75627; CAA53289.1; -.
DR   InterPro; IPR002543; -.
DR   Pfam; PF01580; FtsK_SpoIIIE; 1.
KW   Cell division; ATP-binding; Transmembrane.
FT   TRANSMEM     28     48       POTENTIAL.
FT   TRANSMEM     81    101       POTENTIAL.
FT   TRANSMEM    122    142       POTENTIAL.
FT   TRANSMEM    163    183       POTENTIAL.
FT   NP_BIND     436    443       ATP (POTENTIAL).
SQ   SEQUENCE   778 AA;  85169 MW;  FCE6C15100422061 CRC64;


  Query Match             19.6%;  Score 65.5;  DB 1;  Length 778;
  Best Local Similarity   34.0%;  Pred. No. 1.9;
  Matches   18;  Conservative   12;  Mismatches   18;  Indels    5;  Gaps    3;

Qy       9 LLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHEC---HFPN 58
           :|| ::| :| |: |  :| :|:  |:|: |:|   | | :  |     | ||
Db      15 VLLRRHLRYRLREGC-FILALAFSAFLFIALLSYHRSDPGW-SHSIVVKHVPN 65


RESULT   6
ACH5_CAEEL
ID   ACH5_CAEEL     STANDARD;      PRT;   511 AA.
AC   Q23022; Q17408; P91265; O02559;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   ACETYLCHOLINE RECEPTOR, ALPHA-TYPE SUBUNIT UNC-38 PRECURSOR.
GN   UNC-38 OR F21F3.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=97368239; PubMed=9221782;
RA   Fleming J.T., Squire M.D., Barnes T.M., Tornoe C., Matsuda K.,
RA   Ahnn J., Fire A., Sulston J.E., Barnard E.A., Sattelle D.B.,
RA   Lewis J.A.;
RT   "Caenorhabditis elegans levamisole resistance genes lev-1, unc-29,
RT   and unc-38 encode functional nicotinic acetylcholine receptor
RT   subunits.";
RL   J. Neurosci. 17:5843-5857(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Geisel C., Kramer J., Elliott G.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ACETYLCHOLINE RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X98600; CAA67197.1; -.
DR   EMBL; X98599; CAA67196.1; -.
DR   EMBL; U88175; AAB42282.1; -.
DR   WormPep; F21F3.5; CE09535.
DR   InterPro; IPR001175; -.
DR   InterPro; IPR002394; -.
DR   Pfam; PF00065; neur_chan; 1.
DR   PRINTS; PR00252; NRIONCHANNEL.
DR   PRINTS; PR00254; NICOTINICR.
DR   PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW   Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein;
KW   Transmembrane; Signal.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17    511       ACETYLCHOLINE RECEPTOR, ALPHA-TYPE
FT                                SUBUNIT UNC-38.
FT   DOMAIN       17    261       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    262    282       POTENTIAL.
FT   TRANSMEM    291    311       POTENTIAL.
FT   TRANSMEM    324    344       POTENTIAL.
FT   DOMAIN      345    464       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    465    485       POTENTIAL.
FT   DISULFID    151    165       BY SIMILARITY.
FT   DISULFID    238    239       BY SIMILARITY.
FT   CARBOHYD    124    124       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    248    248       T -> TKHEKLQSLNKKLQPNFARIEKYI (IN
FT                                REF. 2).
SQ   SEQUENCE   511 AA;  59454 MW;  19F690731B23E1A6 CRC64;


  Query Match             18.1%;  Score 60.5;  DB 1;  Length 511;
  Best Local Similarity   27.9%;  Pred. No. 5.1;
  Matches   12;  Conservative   10;  Mismatches   20;  Indels    1;  Gaps    1;

Qy      14 NLTFRRRQTCQLLLEVAW-PLFIFLILISVRLSYPPYEQHECH 55
           |:|   : |     || | |  ||  :  : : : |::: :||
Db     124 NITISTKATLHYTGEVTWEPPAIFKSMCQIDVRWFPFDEQQCH 166


RESULT   7
FD62_SOYBN
ID   FD62_SOYBN     STANDARD;      PRT;   383 AA.
AC   P48631;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   OMEGA-6 FATTY ACID DESATURASE, ENDOPLASMIC RETICULUM ISOZYME 2
DE   (EC 1.14.99.-).
GN   FAD2-2.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Epicotyl;
RX   MEDLINE=96151506; PubMed=8587990;
RA   Heppard E.P., Kinney A.J., Stecca K.L., Miao G.H.;
RT   "Developmental and growth temperature regulation of two different
RT   microsomal omega-6 desaturase genes in soybeans.";
RL   Plant Physiol. 110:311-319(1996).
CC   -!- FUNCTION: ER (MICROSOMAL) OMEGA-6 FATTY ACID DESATURASE INTRODUCES
CC       THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 18:3 FATTY ACIDS,
CC       IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT TO USE
CC       CYTOCHROME B5 AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS
CC       ESTERIFIED TO PHOSPHATIDYLCHOLINE AND, POSSIBLY, OTHER
CC       PHOSPHOLIPIDS.
CC   -!- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM.
CC   -!- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC       AND/ OR BE INVOLVED IN METAL ION BINDING.
CC   -!- SIMILARITY: TO OTHER PLANT ER OMEGA-6 FATTY ACID DESATURASES.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L43921; AAB00860.1; -.
DR   InterPro; IPR001225; -.
DR   Pfam; PF00487; FA_desaturase; 2.
KW   Oxidoreductase; Fatty acid biosynthesis; Endoplasmic reticulum;
KW   Transmembrane.
FT   TRANSMEM      61     81       POTENTIAL.
FT   TRANSMEM      85    105       POTENTIAL.
FT   TRANSMEM     117    137       POTENTIAL.
FT   TRANSMEM     179    199       POTENTIAL.
FT   TRANSMEM     225    245       POTENTIAL.
FT   TRANSMEM     249    269       POTENTIAL.
FT   DOMAIN       105    109       HISTIDINE BOX 1.
FT   DOMAIN       141    145       HISTIDINE BOX 2.
FT   DOMAIN       315    319       HISTIDINE BOX 3.
SQ   SEQUENCE   383 AA;  43967 MW;  F23EF7159B2F9967 CRC64;


  Query Match           17.5%;  Score 58.5;  DB 1;  Length 383;
  Best Local Similarity  28.6%;  Pred. No. 6.8;
  Matches   10;  Conservative   10;  Mismatches   10;  Indels    5;  Gaps    1;

Qy      22 TCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHF 56
           |  : | : |||:     ::: :|  ||::  ||:
Db     181 TLAVTLTLGWPLY-----LALNVSGRPYDRFACHY 210


RESULT   8
SYV_NEUCR
ID   SYV_NEUCR      STANDARD;      PRT;  1093 AA.
AC   P28350;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   VALYL-TRNA SYNTHETASE, MITOCHONDRIAL PRECURSOR (EC 6.1.1.9)
DE   (VALINE--TRNA LIGASE) (VALRS).
GN   CYT-20 OR UN-3.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=74-OR23-1A;
RX   MEDLINE=91304394; PubMed=1830127;
RA   Kubelik A.R., Turcq B., Lambowitz A.M.;
RT   "The Neurospora crassa cyt-20 gene encodes cytosolic and
RT   mitochondrial valyl-tRNA synthetases and may have a second function
RT   in addition to protein synthesis.";
RL   Mol. Cell. Biol. 11:4022-4035(1991).
CC   -!- FUNCTION: MAY HAVE A SECOND FUNCTION IN ADDITION TO PROTEIN
CC       SYNTHESIS.
CC   -!- CATALYTIC ACTIVITY: ATP + L-VALINE + TRNA(VAL) = AMP +
CC       PYROPHOSPHATE + L-VALYL-TRNA(VAL).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL AND CYTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: A SINGLE NUCLEAR GENE PRODUCES BOTH FORMS
CC       BY USE OF ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M64703; -; NOT_ANNOTATED_CDS.
DR   PIR; A41251; A41251.
DR   InterPro; IPR001412; -.
DR   InterPro; IPR002300; -.
DR   InterPro; IPR002303; -.
DR   Pfam; PF00133; tRNA-synt_1; 1.
DR   PRINTS; PR00986; TRNASYNTHVAL.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW   Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW   Mitochondrion; Transit peptide; Alternative initiation.
FT   TRANSIT       1     44       MITOCHONDRION.
FT   CHAIN        45   1093       VALYL-TRNA SYNTHETASE, MITOCHONDRIAL
FT                                ISOFORM.
FT   CHAIN        44   1093       VALYL-TRNA SYNTHETASE, CYTOPLASMIC
FT                                ISOFORM.
FT   INIT_MET     44     44       FOR CYTOPLASMIC ISOFORM.
FT   SITE        179    189       "HIGH" REGION.
FT   SITE        692    696       "KMSKS" REGION.
FT   BINDING     695    695       ATP (BY SIMILARITY).
FT   MUTAGEN     201    201       R->C: GROSS DEFICIENCY OF BOTH MT AND
FT                                CYTOSILIC VALRS ACTIVITIES.
SQ   SEQUENCE   1093 AA;  123352 MW;  3492E40668CAB42C CRC64;


  Query Match           17.1%;  Score 57;  DB 1;  Length 1093;
  Best Local Similarity  26.7%;  Pred. No. 26;
  Matches   16;  Conservative   16;  Mismatches   18;  Indels   10;  Gaps    5;

Qy       4 WPQLRLLLW---KNLTFRRRQTCQLLLEVAWPLFIF----LILISVRLS-YPPYEQHECH 55
           || : :| |   :|| |::      :|| | : |    :|::|::::   |: :  ||
Db     625 WP-MAILGWPNTENLDFKKFFPTS-MLETGWDILFFWVSRMIMLSLKMTGEVPFTEVYCH 682


RESULT   9
CYGF_BOVIN
ID   CYGF_BOVIN     STANDARD;      PRT;  1103 AA.
AC   O02740;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   RETINAL GUANYLYL CYCLASE 2 PRECURSOR (EC 4.6.1.2) (GUANYLATE CYCLASE
DE   2F, RETINAL) (RETGC-2) (ROD OUTER SEGMENT MEMBRANE GUANYLATE CYCLASE
DE   2) (ROS-GC2) (GUANYLATE CYCLASE F) (GC-F).
GN   GUCY2F OR GUC2F.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98238659; PubMed=9571173;
RA   Goraczniak R.M., Duda T., Sharma R.K.;
RT   "Calcium modulated signaling site in type 2 rod outer segment
RT   membrane guanylate cyclase (ROS-GC2).";
RL   Biochem. Biophys. Res. Commun. 245:447-453(1998).
CC   -!- FUNCTION: PROBABLY PLAYS A SPECIFIC FUNCTIONAL ROLE IN THE RODS
CC       AND/OR CONES OF PHOTORECEPTORS. IT MAY BE THE ENZYME INVOLVED IN
CC       THE RESYNTHESIS OF CGMP REQUIRED FOR RECOVERY OF THE DARK STATE
CC       AFTER PHOTOTRANSDUCTION (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
CC   -!- ENZYME REGULATION: ACTIVATED BY GCAP-1; INHIBITED BY CALCIUM.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- PTM: THERE ARE 9 CONSERVED CYSTEINE RESIDUES IN SENSORY GUANYLATE
CC       CYCLASES, 6 IN THE EXTRACELLULAR DOMAIN, WHICH MAY BE INVOLVED IN
CC       INTRA- OR INTERCHAIN DISULFIDE BONDS.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U95958; AAB53864.1; -.
DR   HSSP; Q02846; 1AWL.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001054; -.
DR   InterPro; IPR001828; -.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Lyase; cGMP synthesis; Signal; Transmembrane; Multigene family;
KW   Vision.
FT   SIGNAL        1     46       BY SIMILARITY.
FT   CHAIN        47   1103       RETINAL GUANYLYL CYCLASE 2.
FT   DOMAIN       47    465       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    466    490       POTENTIAL.
FT   DOMAIN      491   1103       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      524    810       PROTEIN KINASE-LIKE.
FT   DOMAIN      813   1064       CATALYTIC.
FT   DISULFID    104    132       BY SIMILARITY.
FT   DISULFID    452    452       INTERCHAIN (PROBABLE).
FT   DISULFID    460    460       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1103 AA;  124261 MW;  EB731E1D8C642AA4 CRC64;


  Query Match             17.1%;  Score 57;  DB 1;  Length 1103;
  Best Local Similarity   34.0%;  Pred. No. 26;
  Matches   18;  Conservative   11;  Mismatches   20;  Indels    4;  Gaps    3;

Qy       1 MACWPQLRLLLW-KNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSY--PPYE 50
           :| ||   |:||   |  :| |      :: | |:: |:|:|: |    |||:
Db       3 LAPWPFSHLMLWFVTLGRQRGQHGLASFKLLWCLWL-LVLMSLPLQVWAPPYK 54


RESULT  10
Y563_HUMAN
ID   Y563_HUMAN     STANDARD;      PRT;   870 AA.
AC   O60309;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN KIAA0563.
GN   KIAA0563.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98290545; PubMed=9628581;
RA   Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA   Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. IX.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:31-39(1998).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB011135; BAA25489.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   870 AA;  94665 MW;  BD76CD70A005FA96 CRC64;


  Query Match             16.9%;  Score 56.5;  DB 1;  Length 870;
  Best Local Similarity   34.4%;  Pred. No. 24;
  Matches   21;  Conservative    6;  Mismatches   19;  Indels   15;  Gaps    4;

Qy       4 WPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIF---LILISVRLSYP-PYEQHECHFPNK 59
           ||   ||:|      |   ||:: | ||      | | |  |  | |:  |  ||| :
Db      22 WP---LLMW--------QLLWLLVKEAQPLEWVKDPLQLTSNPLGPPEPWSSHSSHFPRE 70

Qy      60 A 60
           :
Db      71 S 71


RESULT  11
D5DR_FUGRU
ID   D5DR_FUGRU     STANDARD;      PRT;   463 AA.
AC   P53454;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   D(5)-LIKE DOPAMINE RECEPTOR.
GN   DL.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95309911; PubMed=7789977;
RA   Machae A.D., Brenner S.;
RT   "Analysis of the dopamine receptor family in the compact genome of
RT   the puffer fish Fugu rubripes.";
RL   Genomics 25:436-446(1995).
CC   -!- FUNCTION: RECEPTOR FOR DOPAMINE.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X80177; CAA56457.1; -.
DR   GCRDb; GCR_1105; -.
DR   InterPro; IPR000276; -.
DR   InterPro; IPR000929; -.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00242; DOPAMINER.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     39       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     40     65       1 (POTENTIAL).
FT   DOMAIN       66     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     77    103       2 (POTENTIAL).
FT   DOMAIN      104    112       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    113    135       3 (POTENTIAL).
FT   DOMAIN      136    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    180       4 (POTENTIAL).
FT   DOMAIN      181    198       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    199    223       5 (POTENTIAL).
FT   DOMAIN      224    273       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    274    301       6 (POTENTIAL).
FT   DOMAIN      302    315       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM    316    337       7 (POTENTIAL).
FT   DOMAIN      338    463       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      6      6       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    112    194       BY SIMILARITY.
SQ   SEQUENCE   463 AA;  51095 MW;  7FD627F69A699F6B CRC64;


  Query Match             16.8%;  Score 56;  DB 1;  Length 463;
  Best Local Similarity   28.6%;  Pred. No. 16;
  Matches   14;  Conservative   12;  Mismatches   13;  Indels   10;  Gaps    2;

Qy      17 FRRRQTCQ---LLLEVAWPLFIFLILISVRLSY-------PPYEQHECH 55
           : || | :  .|:: ||| | : :  | |:|::         :|| :|:
Db     147 YERRMTRRFAFLMIAVAWTLSVLISFIPVQLNWHRADNNSSAHEQGDCN 195


RESULT  12
CN2A_BOVIN
ID   CN2A_BOVIN     STANDARD;      PRT;   921 AA.
AC   P14099; Q28064;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CGMP-DEPENDENT 3',5'-CYCLIC PHOSPHODIESTERASE (EC 3.1.4.17)
DE   (CYCLIC GMP STIMULATED PHOSPHODIESTERASE) (CGS-PDE).
GN   PDE2A.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A. (PDE2A1).
RX   MEDLINE=91373395; PubMed=1654333;
RA   Sonnenburg W.K., Mullaney P.J., Beavo J.A.;
RT   "Molecular cloning of a cyclic GMP-stimulated cyclic nucleotide
RT   phosphodiesterase cDNA. Identification and distribution of isozyme
RT   variants.";
RL   J. Biol. Chem. 266:17655-17661(1991).
RN   [2]
RP   SEQUENCE FROM N.A. (PDE2A3).
RC   TISSUE=Brain;
RA   Juilfs D.M., Sonnenburg W.K., Seraji S., Beavo J.A.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 592-921 FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=91104948; PubMed=2176866;
RA   le Trong H., Beier N., Sonnenburg W.K., Stroop S.D., Walsh K.A.,
RA   Beavo J.A., Charbonneau H.;
RT   "Amino acid sequence of the cyclic GMP stimulated cyclic nucleotide
RT   phosphodiesterase from bovine heart.";
RL   Biochemistry 29:10280-10288(1990).
RN   [4]
RP   SEQUENCE OF 613-694 AND 808-868.
RC   TISSUE=Heart;
RX   MEDLINE=87092242; PubMed=3025833;
RA   Charbonneau H., Beier N., Walsh K.A., Beavo J.A.;
RT   "Identification of a conserved domain among cyclic nucleotide
RT   phosphodiesterases from diverse species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9308-9312(1986).
CC   -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC       GUANOSINE 5'-PHOSPHATE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; PDE2A1 (SHOWN
CC       HERE), PDE2A2 AND PDE2A3; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M73512; AAA74559.1; -.
DR   EMBL; L49503; AAA87353.1; -.
DR   PIR; A40981; A40981.
DR   InterPro; IPR002073; -.
DR   InterPro; IPR003018; -.
DR   Pfam; PF01590; GAF; 2.
DR   Pfam; PF00233; PDEase; 1.
DR   PRINTS; PR00387; PDIESTERASE1.
DR   PROSITE; PS00126; PDEASE_I; 1.
KW   Hydrolase; cGMP; Alternative splicing; Membrane; Acetylation.
FT   MOD_RES       1      1       ACETYLATION.
FT   DOMAIN      613    871       CATALYTIC.
FT   VARSPLIC      1     25       MRRQPAASRDLFAQEPVPPGSGDGA -> MGQACGHSILCR
FT                                SQQYPAARPAEPRGQQVFLKPDEPPPPPQPCADS (IN
FT                                ISOFORM PDE2A3).
FT   CONFLICT    204    204       N -> D (IN REF. 2).
FT   CONFLICT    633    633       P -> L (IN REF. 4).
SQ   SEQUENCE   921 AA;  103227 MW;  E29F4C9875E83640 CRC64;


  Query Match             16.8%;  Score 56;  DB 1;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 29;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:     :|   | |:: |      || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT  13
CN2A_HUMAN
ID   CN2A_HUMAN     STANDARD;      PRT;   941 AA.
AC   O00408;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CGMP-DEPENDENT 3',5'-CYCLIC PHOSPHODIESTERASE (EC 3.1.4.17)
DE   (CYCLIC GMP STIMULATED PHOSPHODIESTERASE) (CGS-PDE).
GN   PDE2A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97354299; PubMed=9210593;
RA   Rosman G.J., Martins T.J., Sonnenburg W.K., Beavo J.A., Ferguson K.,
RA   Loughney K.;
RT   "Isolation and characterization of human cDNAs encoding a cGMP-
RT   stimulated 3',5'-cyclic nucleotide phosphodiesterase.";
RL   Gene 191:89-95(1997).
CC   -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC       GUANOSINE 5'-PHOSPHATE.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; PDE2A1, PDE2A2 AND
CC       PDE2A3 (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; U67733; AAC51320.1; -.
DR   MIM; 602658; -.
DR   InterPro; IPR002073; -.
DR   InterPro; IPR003018; -.
DR   Pfam; PF01590; GAF; 2.
DR   Pfam; PF00233; PDEase; 1.
DR   PRINTS; PR00387; PDIESTERASE1.
DR   PROSITE; PS00126; PDEASE_I; 1.
KW   Hydrolase; cGMP; Alternative splicing; Membrane.
SQ   SEQUENCE   941 AA;  105716 MW;  9797609B487FD64E CRC64;


  Query Match             16.8%;  Score 56;  DB 1;  Length 941;
  Best Local Similarity   37.0%;  Pred. No. 30;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy      3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
          |:     :|    | |:: |     || ||:||  ||  |  :|| |
Db    369 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 414


RESULT  14
ACHX_ONCVO
ID   ACHX_ONCVO     STANDARD;      PRT;   436 AA.
AC   P54247;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   ACETYLCHOLINE RECEPTOR PROTEIN, NON-ALPHA CHAIN (FRAGMENT).
OS   Onchocerca volvulus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC   Onchocercidae; Onchocerca.
OX   NCBI_TaxID=6282;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94299155; PubMed=8026747;
RA   Ajuh P.M., Egwang T.G.;
RT   "Cloning of a cDNA encoding a putative nicotinic acetylcholine
RT   receptor subunit of the human filarial parasite Onchocerca
RT   volvulus.";
RL   Gene 144:127-129(1994).
CC   -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC       EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC       LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC       MEMBRANE.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L20465; AAA21823.1; -.
DR   EMBL; L12543; AAA29415.1; -.
DR   InterPro; IPR001175; -.
DR   Pfam; PF00065; neur_chan; 1.
DR   PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW   Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
KW   Multigene family.
FT   NON_TER       1      1
FT   DOMAIN       <1    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    219       POTENTIAL.
FT   TRANSMEM    227    245       POTENTIAL.
FT   TRANSMEM    261    280       POTENTIAL.
FT   DOMAIN      281    404       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    405    423       POTENTIAL.
FT   DISULFID     89    103       BY SIMILARITY.
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    140    140       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   436 AA;  51340 MW;  58051BFE59A4555F CRC64;


  Query Match             16.6%;  Score 55.5;  DB 1;  Length 436;
  Best Local Similarity   27.3%;  Pred. No. 17;
  Matches   12;  Conservative   10;  Mismatches   21;  Indels    1;  Gaps    1;

Qy     14 NLTFRRRQTCQLLLEVAWPL-FIFLILISVRLSYPPYEQHECHF 56
          |:|   : | :   :| |    || |  : : : |::: |||
Db     62 NITISTKATLRYDGQVTWDSPAIFKTLCQIDVRWFPFDEQNCHF 105


RESULT  15
ATS4_HUMAN
ID   ATS4_HUMAN     STANDARD;      PRT;   837 AA.
AC   O75173; Q9UN83;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ADAM-TS 4 PRECURSOR (EC 3.4.24.-) (A DISINTEGRIN AND METALLOPROTEINASE
DE   WITH THROMBOSPONDIN MOTIFS 4) (ADAMTS-4) (ADAM-TS4) (AGGRECANASE 1)
DE   (ADMP-1).
GN   ADAMTS4 OR KIAA0688.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA   Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99286303; PubMed=10356395;
RA   Tortorella M.D., Burn T.C., Pratta M.A., Abbaszade I., Hollis J.M.,
RA   Liu R., Rosenfeld S.A., Copeland R.A., Decicco C.P., Wynn R.,
RA   Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,
RA   Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K.,
RA   Hillman M.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,
RA   Trzaskos J.M., Arner E.C.;
RT   "Purification and cloning of aggrecanase-1: a member of the ADAMTS
RT   family of proteins.";
RL   Science 284:1664-1666(1999).
RN   [3]
RP   PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX   MEDLINE=20400518; PubMed=10827174;
RA   Tortorella M., Pratta M., Liu R.Q., Abbaszade I., Ross H., Burn T.,
RA   Arner E.;
RT   "The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for
RT   aggrecan substrate recognition and cleavage.";
RL   J. Biol. Chem. 275:25791-25797(2000).
CC   -!- FUNCTION: CLEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC       INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
CC       DESTRUCTION OF AGGRECAN IN ARTHRITIC DISEASES.
CC   -!- CATALYTIC ACTIVITY: CLEAVES AGGRECAN AT THE 392-GLU-|-ALA-393
CC       SITE.
CC   -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC       MATRIX (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG AND HEART. EXPRESSED
CC       AT VERY LOW LEVEL IN PLACENTA AND SKELETAL MUSCLES.
CC   -!- INDUCTION: BY INTERLEUKIN 1.
CC   -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT
CC       FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B (ZINC
CC       METALLOPROTEASE); ALSO KNOWN AS THE REPROLYSIN SUBFAMILY.
```

```
CC   -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC   -!- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB014588; BAA31663.1; -.
DR   EMBL; AF148213; AAD41494.1; -.
DR   MIM; 603876; -.
DR   HSSP; P34179; 1IAG.
DR   MEROPS; M12.221; -.
DR   InterPro; IPR000130; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001590; -.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF01421; Reprolysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS50092; TSP1; 1.
DR   PROSITE; PS00427; DISINTEGRINS; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Extracellular matrix.
FT   SIGNAL       1     51       POTENTIAL.
FT   PROPEP      52    212
FT   CHAIN      213    837       ADAM-TS 4.
FT   SITE       194    194       CYSTEINE SWITCH (POTENTIAL).
FT   METAL      361    361       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE   362    362       BY SIMILARITY.
FT   METAL      365    365       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL      371    371       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN     437    519       DISINTEGRIN-LIKE.
FT   DOMAIN     520    576       TSP-TYPE 1 1.
FT   DOMAIN     577    685       CYS-RICH.
FT   DOMAIN     686    837       SPACER.
FT   DOMAIN     247    252       POLY-ALA.
FT   CARBOHYD    68     68       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    77     77       A -> T (IN REF. 1).
SQ   SEQUENCE   837 AA;  90224 MW;  5DF9C9AC137DF41F CRC64;


  Query Match             16.6%;  Score 55.5;  DB 1;  Length 837;
  Best Local Similarity   32.6%;  Pred. No. 31;
  Matches   15;  Conservative   10;  Mismatches   14;  Indels    7;  Gaps    2;

Qy      21 QTCQLL--LEVAWPLFIFLILI-----SVRLSYPPYEQHECHFPNK 59
           | | ||  : ::| ::: |:|:     | ||: |   : |  || |
Db      21 QPCLLLPIVPLSWLVWLLLLLLASLLPSARLASPLPREEEIVFPEK 66


Search completed: May 31, 2001, 13:08:58
Job time: 299 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 31, 2001, 12:19:49 ; Search time 39.12 Seconds
(without alignments)
105.403 Million cell updates/sec

Title: US-09-526-193A-1_COPY_1_60
Perfect score: 334
Sequence: 1 MACWPQLRLLLWKNLTFRRR.....SVRLSYPPYEQHECHFPNKA 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	86	25.7	1704	2	S71363	probable ATP-bindi
2	84	25.1	1704	2	A59188	ATP-binding casset
3	71.5	21.4	707	2	T09340	hypothetical prote
4	65.5	19.6	778	2	S43132	sporulation protei
5	60.5	18.1	511	2	T43634	nicotinic acetylch
6	60.5	18.1	534	2	T25720	hypothetical prote
7	60.5	18.1	598	2	T05130	hypothetical prote
8	60	18.0	1802	2	T33783	hypothetical prote
9	59.5	17.8	466	2	PC4296	nicotinic acetylch
10	58.5	17.5	383	2	T07688	omega-6 desaturase
11	58	17.4	250	2	D75317	hypothetical prote
12	57	17.1	1093	2	A41251	valine--tRNA ligas
13	57	17.1	1103	2	JC5581	guanylate cyclase
14	56.5	16.9	1400	2	T22644	hypothetical prote
15	56.5	16.9	1977	2	S54771	sodium channel alp
16	56.5	16.9	3268	2	S69625	hypothetical prote
17	56	16.8	373	2	JC7289	G-protein coupled
18	56	16.8	463	2	B56849	dopamine receptor-
19	56	16.8	512	2	E81938	probable apolipopr
20	56	16.8	524	2	H81166	apolipoprotein N-a
21	56	16.8	830	2	T04848	protein kinase hom
22	56	16.8	921	1	A40981	3',5'-cyclic-nucle
23	56	16.8	1215	2	B72029	helicase, Snf2/Rad
24	56	16.8	2201	2	A54774	ATP binding casset
25	55.5	16.6	299	2	F83301	conserved hypothet
26	55.5	16.6	372	2	T20289	hypothetical prote
27	55.5	16.6	378	2	T14269	Delta12 fatty acid
28	55.5	16.6	495	2	G70852	hypothetical prote
29	55.5	16.6	706	2	T20052	hypothetical prote
30	55.5	16.6	837	2	T00355	hypothetical prote
31	55	16.5	293	2	T32229	hypothetical prote
32	55	16.5	357	2	H64855	probable membrane
33	55	16.5	928	1	JC2486	3',5'-cyclic-nucle
34	55	16.5	1976	2	I56555	sodium channel pro
35	54.5	16.3	383	2	T10480	Delta12 fatty acid
36	54.5	16.3	454	2	T32974	hypothetical prote
37	54	16.2	401	2	B83369	conserved hypothet
38	54	16.2	442	2	A83743	magnesium citrate
39	54	16.2	547	2	S61032	hypothetical prote
40	54	16.2	825	2	T27852	hypothetical prote
41	54	16.2	2005	2	B25019	sodium channel pro
42	54	16.2	2005	2	A46269	sodium channel alp
43	53.5	16.0	333	2	I65989	G protein-coupled
44	53.5	16.0	538	2	D82180	probable sensor ki
45	53.5	16.0	778	2	T05341	S-receptor kinase

ALIGNMENTS

RESULT 1
S71363
probable ATP-binding cassette transporter ABC-3 - human
N;Alternate names: ATP-binding cassette transporter ABC-C
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 17-Mar-2000
C;Accession: S71363
R;Klugbauer, N.; Hofmann, F.
FEBS Lett. 391, 61-65, 1996
A;Title: Primary structure of a novel ABC transporter with a chromosomal localization
A;Reference number: S71363; MUID:96326608
A;Accession: S71363
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1704 <KLU>
A;Cross-references: EMBL:X97187; NID:g1514529; PIDN:CAA65825.1; PID:e243436; PID:g151
A;Experimental source: cell line medullary thyroid carcinoma
C;Genetics:
A;Gene: GDB:ABC3
A;Cross-references: GDB:3770735; OMIM:601615
A;Map position: 16p13.3-16p13.3
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
C;Keywords: ATP binding; P-loop; phosphoprotein; transmembrane protein
F;255-283/Domain: transmembrane #status predicted <TM1>
F;307-329/Domain: transmembrane #status predicted <TM2>
F;345-364/Domain: transmembrane #status predicted <TM3>
F;373-394/Domain: transmembrane #status predicted <TM4>
F;401-422/Domain: transmembrane #status predicted <TM5>
F;452-475/Domain: transmembrane #status predicted <TM6>
F;549-739/Domain: ATP-binding cassette homology <ABC1>
F;566-573/Region: nucleotide-binding motif A (P-loop)
F;685-690/Region: nucleotide-binding motif B
F;1100-1120/Domain: transmembrane #status predicted <TM7>
F;1145-1169/Domain: transmembrane #status predicted <TM8>
F;1181-1207/Domain: transmembrane #status predicted <TM9>
F;1215-1236/Domain: transmembrane #status predicted <TM10>
F;1245-1264/Domain: transmembrane #status predicted <TM11>
F;1299-1324/Domain: transmembrane #status predicted <TM12>
F;1399-1590/Domain: ATP-binding cassette homology <ABC2>
F;1416-1423/Region: nucleotide-binding motif A (P-loop)
F;1535-1540/Region: nucleotide-binding motif B
F;674,866,1524/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #s
F;1344/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status pr

Query Match 25.7%; Score 86; DB 2; Length 1704;
Best Local Similarity 47.7%; Pred. No. 0.016;
Matches 21; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

```
Qy       1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL 44
           ||   || |||||| | ::|:    :||:  ||    ||| :||
Db       1 MAVLRQLALLLWKNYTLQKRKVLVTVLELFLPLLFSGILIWLRL 44
```

```
RESULT   2
A59188
ATP-binding cassette transporter ABC3 - human
C;Species: Homo sapiens (man)
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 04-Mar-2000
C;Accession: A59188
R;Connors, T.D.; van Raay, T.J.; Petry, L.R.; Klinger, K.W.; Landes, G.M.; Burn, T.C.
Genomics 39, 231-234, 1997
A;Title: The cloning of a human ABC gene (ABC3) mapping to chromosome 16p13.3.
A;Reference number: A59188; MUID:97179225
A;Accession: A59188
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1704 <CON>
A;Cross-references: GB:U78735; NID:g1699037; PIDN:AAC50967.1; PID:g1699038
C;Genetics:
A;Gene: GDB:ABC3
A;Cross-references: GDB:3770735; OMIM:601615
A;Map position: 16p13.3-16p13.3
C;Superfamily: ATP-binding cassette homology


  Query Match             25.1%;  Score 84;  DB 2;  Length 1704;
  Best Local Similarity   47.7%;  Pred. No. 0.028;
  Matches   21;  Conservative    6;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL 44
           ||   || ||||||  | ::|:    :||:  ||    ||| :||
Db       1 MAVLRQLALLLWKNYTLQKRKVLVTVLELFLPLLFPGILIWLRL 44


RESULT   3
T09340
hypothetical protein T26M18.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C;Accession: T09340
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16650
A;Accession: T09340
A;Molecule type: DNA
A;Residues: 1-707 <BEV>
A;Cross-references: EMBL:AL078606; GSPDB:GN00062; ATSP:T26M18.20
A;Experimental source: cultivar Columbia; BAC clone T26M18
C;Genetics:
A;Gene: ATSP:T26M18.20
A;Map position: 4
A;Introns: 18/2; 58/3; 162/3; 197/3; 354/2; 455/1; 520/3; 584/2; 629/1


  Query Match             21.4%;  Score 71.5;  DB 2;  Length 707;
  Best Local Similarity   36.7%;  Pred. No. 0.43;
  Matches   22;  Conservative    7;  Mismatches   24;  Indels    7;  Gaps    2;

Qy       1 MACWPQLRLLLW-----KNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECH 55
           ||| | |  ||      ||:|| :      :: ||| |::  : || |   |  |  | |
Db     390 MACGPALAGLLQTDFKIKNVTFNQDTLPGWVMAVAWLLYLVWLAISFR--EPAREPEEIH 447


RESULT   4
S43132
sporulation protein spoIIIE - Coxiella burnetii
C;Species: Coxiella burnetii
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 08-Oct-1999
C;Accession: S43132; S31759
R;Oswald, W.
submitted to the EMBL Data Library, November 1993
A;Reference number: S43131
A;Accession: S43132
A;Molecule type: DNA
A;Residues: 1-778 <OSW>
A;Cross-references: EMBL:X75627; NID:g468527; PIDN:CAA53289.1; PID:g468529
C;Genetics:
A;Gene: spoIIIE


  Query Match             19.6%;  Score 65.5;  DB 2;  Length 778;
  Best Local Similarity   34.0%;  Pred. No. 2.5;
  Matches   18;  Conservative   12;  Mismatches   18;  Indels    5;  Gaps    3;

Qy       9 LLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHEC---HFPN 58
           :|| ::| :| |: |  :| :|:  |:|: |:|   | | :  |     | ||
Db      15 VLLRRHLRYRLREGC-FILALAFSAFLFIALLSYHRSDPGW-SHSIVVKHVPN 65


RESULT   5
T43634
nicotinic acetylcholine receptor alpha chain - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 11-May-2000
C;Accession: T43634
R;Fleming, J.T.; Squire, M.D.; Barnes, T.M.; Tornoe, C.; Matsuda, K.; Ahnn, J.; Fire,
J. Neurosci. 17, 5843-5857, 1997
A;Title: Caenorhabditis elegans levamisole resistance genes lev-1, unc-29 and unc-38
A;Reference number: Z22588; MUID:97368239
A;Accession: T43634
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-511 <FLE>
A;Cross-references: EMBL:X98599; PIDN:CAA67196.1
C;Genetics:
A;Introns: 43/2; 66/3; 118/2; 168/3; 248/3; 341/1; 484/3
A;Note: unc-38
C;Superfamily: acetylcholine receptor


  Query Match             18.1%;  Score 60.5;  DB 2;  Length 511;
  Best Local Similarity   27.9%;  Pred. No. 7.1;
  Matches   12;  Conservative   10;  Mismatches   20;  Indels    1;  Gaps    1;

Qy      14 NLTFRRRQTCQLLLEVAW-PLFIFLILISVRLSYPPYEQHECH 55
           |:|   : |     || | |  ||  :  : : : |::: :||
Db     124 NITISTKATLHYTGEVTWEPPAIFKSMCQIDVRWFPFDEQQCH 166


RESULT   6
T25720
hypothetical protein F21F3.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T25720
R;Geisel, C.; Kramer, J.; Elliott, G.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid F21F3.
A;Reference number: Z20075
A;Accession: T25720
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-534 <GEI>
A;Cross-references: EMBL:U88175; PIDN:AAB42282.1; GSPDB:GN00019; CESP:F21F3.5
A;Experimental source: strain Bristol N2; clone F21F3
C;Genetics:
A;Gene: CESP:F21F3.5
A;Map position: 1
A;Introns: 43/2; 66/3; 118/2; 168/3; 248/3; 271/3; 364/1; 507/3
C;Superfamily: acetylcholine receptor


  Query Match             18.1%;  Score 60.5;  DB 2;  Length 534;
  Best Local Similarity   27.9%;  Pred. No. 7.4;
  Matches   12;  Conservative   10;  Mismatches   20;  Indels    1;  Gaps    1;

Qy      14 NLTFRRRQTCQLLLEVAW-PLFIFLILISVRLSYPPYEQHECH 55
```

```
          |:|   : |     || | |  ||  :  : : : |::: :||
Db    124 NITISTKATLHYTGEVTWEPPAIFKSMCQIDVRWFPFDEQQCH 166


RESULT   7
T05130
hypothetical protein F7H19.170 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C;Accession: T05130
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, July 1998
A;Reference number: Z15399
A;Accession: T05130
A;Molecule type: DNA
A;Residues: 1-598 <BEV>
A;Cross-references: EMBL:AL031018
A;Experimental source: cultivar Columbia; BAC clone F7H19
C;Genetics:
A;Map position: 4
A;Introns: 18/2; 58/3; 162/3; 351/2; 454/1; 515/3; 579/2
A;Note: F7H19.170


  Query Match             18.1%;  Score 60.5;  DB 2;  Length 598;
  Best Local Similarity   30.5%;  Pred. No. 8.2;
  Matches   18;  Conservative   11;  Mismatches   21;  Indels    9;  Gaps    3;

Qy      1 MACWP------QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHE 53
          |||  |      |:|  ::| ||| :      :: |||   |:|: :::    |  |  |
Db    387 MACGPALAGLLQIRFKIYK-LTFNQDTLPGWVMAVAW--LIYLVWLAISFREPAREPEE 442


RESULT   8
T33783
hypothetical protein Y39D8C.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C;Accession: T33783
R;Becker, M.; Graves, T.; Yoakum, M.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of C. elegans cosmid Y39D8C.
A;Reference number: Z21408
A;Accession: T33783
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1802 <BEC>
A;Cross-references: EMBL:AF101313; PIDN:AAC69223.1; GSPDB:GN00023; CESP:Y39D8C.1
A;Experimental source: strain Bristol N2; clone Y39D8C
C;Genetics:
A;Gene: CESP:Y39D8C.1
A;Map position: 5
A;Introns: 45/3; 114/1; 195/1; 230/3; 543/3; 794/1; 849/1; 1036/2; 1099/1; 1132/3; 1165/
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology


  Query Match             18.0%;  Score 60;  DB 2;  Length 1802;
  Best Local Similarity   42.9%;  Pred. No. 26;
  Matches   15;  Conservative    3;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFI 35
          |    |||||||||   : |     | |:  || :
Db      1 MGALGQLRLLLWKNWLQQLRSPWFALFELVVPLIL 35


RESULT   9
PC4296
nicotinic acetylcholine receptor alpha chain - nematode (Trichostrongylus colubriformis)
N;Alternate names: tar-1 protein
C;Species: Trichostrongylus colubriformis
C;Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 21-Aug-1998
C;Accession: PC4296
R;Wiley, L.J.; Weiss, A.S.; Sangster, N.C.; Li, Q.
Gene 182, 97-100, 1996
A;Title: Cloning and sequence analysis of the candidate nicotinic acetylcholine recep
A;Reference number: PC4296; MUID:97136696
A;Accession: PC4296
A;Molecule type: DNA
A;Residues: 1-466 <WIL>
A;Cross-references: GB:U56903
C;Genetics:
A;Gene: tar-1
C;Superfamily: acetylcholine receptor
F;214-239/Domain: transmembrane #status predicted <TM1>
F;247-270/Domain: transmembrane #status predicted <TM2>
F;276-299/Domain: transmembrane #status predicted <TM3>
F;415-436/Domain: transmembrane #status predicted <TM4>


  Query Match             17.8%;  Score 59.5;  DB 2;  Length 466;
  Best Local Similarity   27.9%;  Pred. No. 8.6;
  Matches   12;  Conservative   10;  Mismatches   20;  Indels    1;  Gaps    1;

Qy     14 NLTFRRRQTCQLLLEVAW-PLFIFLILISVRLSYPPYEQHECH 55
          |:|   : |     || | |  ||  :  : : : |::: :||
Db     81 NITISTKATLHYSGEVTWEPPAIFKSMCQIDVRWFPFDEQQCH 123


RESULT  10
T07688
omega-6 desaturase FAD2-2, microsomal - soybean
C;Species: Glycine max (soybean)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000
C;Accession: T07688
R;Heppard, E.P.; Kinney, A.J.; Stecca, K.L.; Miao, G.H.
Plant Physiol. 110, 311-319, 1996
A;Title: Developmental and growth temperature regulation of two different microsomal
A;Reference number: Z16095; MUID:96151506
A;Accession: T07688
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-383 <HEP>
A;Cross-references: EMBL:L43921; NID:g904153; PIDN:AAB00860.1; PID:g904154
A;Experimental source: epicotyl
C;Genetics:
A;Gene: FAD2-2
C;Function:
A;Description: involved in production of polyunsaturated lipids
A;Note: constitutively expressed in both vegetative tissues and developing seeds
C;Superfamily: omega-3 fatty acid desaturase


  Query Match             17.5%;  Score 58.5;  DB 2;  Length 383;
  Best Local Similarity   28.6%;  Pred. No. 9.6;
  Matches   10;  Conservative   10;  Mismatches   10;  Indels    5;  Gaps    1;

Qy     22 TCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHF 56
          |  : | : |||:     ::: :|  ||::  ||:
Db    181 TLAVTLTLGWPLY-----LALNVSGRPYDRFACHY 210


RESULT  11
D75317
hypothetical protein DR2083 - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: D75317
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896
A;Accession: D75317
A;Status: preliminary
A;Molecule type: DNA
```

```
A;Residues: 1-250 <WHI>
A;Cross-references: GB:AE002044; GB:AE000513; NID:g6459872; PIDN:AAF11637.1; PID:g645988
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2083
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR2083


  Query Match             17.4%;  Score 58;  DB 2;  Length 250;
  Best Local Similarity   38.9%;  Pred. No. 7.5;
  Matches   21;  Conservative    5;  Mismatches   14;  Indels   14;  Gaps    4;

Qy       5 PQLRLLLWKN---LTFRRRQTCQLLLEVAWPLFIFL--------ILIS-VRLSY 46
           |:  |::||'     ||| |     | |  ||| : |        :|||  |||| 
Db     105 PETLLIVWKKGDAATFRPRGV--PLRENIWPLLLLLPLSLAVGSLLISRPRLSY 156


RESULT  12
A41251
valine--tRNA ligase (EC 6.1.1.9) - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 14-Nov-1997
C;Accession: A41251
R;Kubelik, A.R.; Turcq, B.; Lambowitz, A.M.
Mol. Cell. Biol. 11, 4022-4035, 1991
A;Title: The Neurospora crassa cyt-20 gene encodes cytosolic and mitochondrial valyl-tRN
A;Reference number: A41251; MUID:91304394
A;Accession: A41251
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1093 <KUB>
A;Cross-references: GB:M64703
C;Superfamily: valine--tRNA ligase
C;Keywords: aminoacyl-tRNA synthetase; ligase; mitochondrion; protein biosynthesis


  Query Match             17.1%;  Score 57;  DB 2;  Length 1093;
  Best Local Similarity   26.7%;  Pred. No. 38;
  Matches   16;  Conservative   16;  Mismatches   18;  Indels   10;  Gaps    5;

Qy       4 WPQLRLLLW---KNLTFRRRQTCQLLLEVAWPLFIF----LILISVRLS-YPPYEQHECH 55
           || : :| |   :|| |::      :||  | :  |    :|::|::::   |: :  ||
Db     625 WP-MAILGWPNTENLDFKKFFPTS-MLETGWDILFFWVSRMIMLSLKMTGEVPFTEVYCH 682


RESULT  13
JC5581
guanylate cyclase (EC 4.6.1.2) ROS-GC2 precursor - bovine
N;Alternate names: guanyl cyclase; guanylyl cyclase
C;Species: Bos primigenius taurus (cattle)
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 16-Jul-1999
C;Accession: JC5581
R;Goraczniak, R.; Duda, T.; Sharma, R.K.
Biochem. Biophys. Res. Commun. 234, 666-670, 1997
A;Title: Structural and functional characterization of a second subfamily member of the
A;Reference number: JC5581; MUID:97318835
A;Accession: JC5581
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1103 <GOR>
A;Cross-references: GB:U95958; NID:g2072999; PIDN:AAB53864.1; PID:g2073000
A;Experimental source: retina
C;Comment: This enzyme belongs to the subfamily of calcium-modulated rod outer segment m
C;Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain homo
C;Keywords: phosphorus-oxygen lyase
F;1-50/Domain: signal sequence #status predicted <SIG>
F;51-465/Domain: extracellular #status predicted <EXT>
F;466-490/Domain: transmembrane #status predicted <TRM>
F;523-816/Domain: protein kinase homology <KIN>
F;836-1064/Domain: guanylate cyclase catalytic domain homology <GCC>
F;1099-1103/Region: signature


  Query Match             17.1%;  Score 57;  DB 2;  Length 1103;
  Best Local Similarity   34.0%;  Pred. No. 39;
  Matches   18;  Conservative   11;  Mismatches   20;  Indels    4;  Gaps    3;

Qy       1 MACWPQLRLLLW-KNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSY--PPYE 50
           :| ||   |:||   |  :| |      :: | |:: |:|:|: |   |||: 
Db       3 LAPWPFSHLMLWFVTLGRQRGQHGLASFKLLWCLWL-LVLMSLPLQVWAPPYK 54


RESULT  14
T22644
hypothetical protein F54D1.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22644
R;Lennard, N.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19592
A;Accession: T22644
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1400 <WIL>
A;Cross-references: EMBL:Z77132; PIDN:CAB00861.1; GSPDB:GN00022; CESP:F54D1.5
A;Experimental source: clone F54D1
C;Genetics:
A;Gene: CESP:F54D1.5
A;Map position: 4
A;Introns: 21/2; 51/2; 205/2; 276/3; 364/2; 394/2; 466/3; 507/3; 536/3; 599/3; 672/2;


  Query Match             16.9%;  Score 56.5;  DB 2;  Length 1400;
  Best Local Similarity   40.0%;  Pred. No. 55;
  Matches   14;  Conservative    4;  Mismatches   10;  Indels    7;  Gaps    1;

Qy      15 LTFRRRQTCQLLLEVA-------WPLFIFLILISV 42
           | |   ||| |||| :       |  ||: : :||
Db     902 LIFLTTQTCILLLETSLKPSKYEWITFIYTVTLSV 936


RESULT  15
S54771
sodium channel alpha subunit - human
C;Species: Homo sapiens (man)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C;Accession: S54771
R;Klugbauer, N.; Lacinova, L.; Flockerzi, V.; Hofmann, F.
EMBO J. 14, 1084-1090, 1995
A;Title: Structure and functional expression of a new member of the tetrodotoxin-sens
A;Reference number: S54771; MUID:95237189
A;Accession: S54771
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1977 <KLU>
A;Cross-references: EMBL:X82835; NID:g758109; PIDN:CAA58042.1; PID:g758110
C;Superfamily: sodium channel protein
C;Keywords: duplication


  Query Match             16.9%;  Score 56.5;  DB 2;  Length 1977;
  Best Local Similarity   22.4%;  Pred. No. 76;
  Matches   11;  Conservative   19;  Mismatches   18;  Indels    1;  Gaps    1;

Qy      10 LLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPN 58
           | ::::   |::| :::|| |  :| :: :: : | :  |  :: :| |
Db    1198 LAFEDIYIERKKTIKIILEYADKIFTYIFILEMLLKWIAY-GYKTYFTN 1245


Search completed: May 31, 2001, 13:05:24
Job time: 2735 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 31, 2001, 13:03:19 ; Search time 54.28 Seconds
(without alignments)
14.331 Million cell updates/sec

Title: US-09-526-193A-1_COPY_1_60
Perfect score: 334
Sequence: 1 MACWPQLRLLLWKNLTFRRR.....SVRLSYPPYEQHECHFPNKA 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 280389 seqs, 12964817 residues

Total number of hits satisfying chosen parameters: 280389

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	334	100.0	2261	1	PCT-US01-04098A-1212	Sequence 1212, Ap
2	334	100.0	2261	5	US-09-526-193A-1	Sequence 1, Appli
3	334	100.0	2263	1	PCT-US01-04098A-3180	Sequence 3180, Ap
4	205	61.4	302	1	PCT-US01-11988-908	Sequence 908, App
5	205	61.4	302	5	US-09-833-245-908	Sequence 908, App
6	117	35.0	21	5	US-09-526-193A-37	Sequence 37, Appl
7	62.5	18.7	774	6	US-60-248-505-1176	Sequence 1176, Ap
8	56.5	16.9	551	1	PCT-US01-04098A-1326	Sequence 1326, Ap
9	56.5	16.9	595	1	PCT-US01-04098A-3294	Sequence 3294, Ap
10	56	16.8	284	1	PCT-US01-01310-71	Sequence 71, Appl
11	56	16.8	378	1	PCT-US01-01310-77	Sequence 77, Appl
12	53	15.9	183	1	PCT-US01-01310-102	Sequence 102, App
13	53	15.9	183	1	PCT-US01-01332-806	Sequence 806, App
14	53	15.9	370	5	US-09-383-745-1	Sequence 1, Appli
15	53	15.9	379	1	PCT-US01-01310-79	Sequence 79, Appl
16	53	15.9	586	6	US-60-248-505-1178	Sequence 1178, Ap
17	52	15.6	191	5	US-09-811-284-161	Sequence 161, App
18	52	15.6	695	6	US-60-248-505-696	Sequence 696, App
19	51.5	15.4	333	5	US-09-826-509-509	Sequence 509, App
20	51	15.3	269	5	US-09-383-745-3	Sequence 3, Appli
21	50.5	15.1	446	5	US-09-826-509-487	Sequence 487, App
22	50.5	15.1	1422	4	US-08-467-344A-81	Sequence 81, Appl
23	50	15.0	150	1	PCT-US01-03782A-115	Sequence 115, App
24	50	15.0	366	1	PCT-US01-11988-2066	Sequence 2066, Ap
25	50	15.0	366	5	US-09-781-417-46	Sequence 46, Appl
26	50	15.0	366	5	US-09-833-245-2066	Sequence 2066, Ap
27	49.5	14.8	319	1	PCT-US01-04098A-1067	Sequence 1067, Ap
28	49.5	14.8	354	1	PCT-US01-04098A-3035	Sequence 3035, Ap
29	49.5	14.8	1052	1	PCT-US01-04098A-1000	Sequence 1000, Ap
30	49	14.7	127	1	PCT-US01-01321-1800	Sequence 1800, Ap
31	49	14.7	180	5	US-09-739-449-11288	Sequence 11288, A
32	49	14.7	803	1	PCT-US01-04098A-2968	Sequence 2968, Ap
33	48	14.4	225	1	PCT-US01-11988-1411	Sequence 1411, Ap
34	48	14.4	225	1	PCT-US01-11988-1413	Sequence 1413, Ap
35	48	14.4	225	5	US-09-833-245-1411	Sequence 1411, Ap
36	48	14.4	225	5	US-09-833-245-1413	Sequence 1413, Ap
37	48	14.4	283	1	PCT-US01-11988-821	Sequence 821, App
38	48	14.4	283	5	US-09-833-245-821	Sequence 821, App
39	48	14.4	696	1	PCT-US01-01312-959	Sequence 959, App
40	47	14.1	172	1	PCT-US01-11988-1412	Sequence 1412, Ap
41	47	14.1	172	5	US-09-833-245-1412	Sequence 1412, Ap
42	47	14.1	618	5	US-09-739-449-10498	Sequence 10498, A
43	47	14.1	766	1	PCT-US01-04098A-1921	Sequence 1921, Ap
44	47	14.1	1416	5	US-09-836-712-2	Sequence 2, Appli
45	46.5	13.9	130	5	US-09-836-377-298	Sequence 298, App

ALIGNMENTS

```
RESULT   1
PCT-US01-04098A-1212
; Sequence 1212, Application PC/TUS0104098A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-029
;  CURRENT APPLICATION NUMBER: PCT/US01/04098A
;  CURRENT FILING DATE:  2001-02-05
;  PRIOR APPLICATION NUMBER: Not Yet Assigned
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: 09/728,422
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 09/693,325
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/663,561
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 09/654,936
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 09/620,325
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/598,075
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 3960
;  SOFTWARE: Custom
; SEQ ID NO 1212
;   LENGTH: 2261
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-04098A-1212


  Query Match            100.0%;  Score 334;  DB 1;  Length 2261;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-30;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   2
US-09-526-193A-1
; Sequence 1, Application US/09526193A
; GENERAL INFORMATION:
;  APPLICANT: Hayden, Michael R.
```

```
;  APPLICANT:  Brooks-Wilson, Angela R.
;  APPLICANT:  Pimstone, Simon N.
;  TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
;  TITLE OF INVENTION:  CHOLESTEROL LEVELS
;  FILE REFERENCE: 50110/002005
;  CURRENT APPLICATION NUMBER: US/09/526,193A
;  CURRENT FILING DATE:  2000-03-15
;  PRIOR APPLICATION NUMBER: 60/124,702
;  PRIOR FILING DATE: 1999-03-15
;  PRIOR APPLICATION NUMBER: 60/138,048
;  PRIOR FILING DATE: 1999-06-08
;  PRIOR APPLICATION NUMBER: 60/139,600
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: 60/151,977
;  PRIOR FILING DATE: 1999-09-01
;  NUMBER OF SEQ ID NOS: 287
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2261
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-526-193A-1


  Query Match             100.0%;  Score 334;  DB 5;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-30;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   3
PCT-US01-04098A-3180
; Sequence 3180, Application PC/TUS0104098A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-029
;  CURRENT APPLICATION NUMBER: PCT/US01/04098A
;  CURRENT FILING DATE:  2001-02-05
;  PRIOR APPLICATION NUMBER: Not Yet Assigned
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: 09/728,422
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 09/693,325
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/663,561
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 09/654,936
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 09/620,325
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/598,075
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 3960
;  SOFTWARE: Custom
; SEQ ID NO 3180
;   LENGTH: 2263
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-04098A-3180


  Query Match             100.0%;  Score 334;  DB 1;  Length 2263;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-30;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 62


RESULT   4
PCT-US01-11988-908
; Sequence 908, Application PC/TUS0111988
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF546PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/11988
;  CURRENT FILING DATE:  2001-01-12
;  PRIOR APPLICATION NUMBER: 60/229, 358
;  PRIOR FILING DATE: 2000-04-12
;  PRIOR APPLICATION NUMBER: 60/256, 931
;  PRIOR FILING DATE: 2000-12-21
;  PRIOR APPLICATION NUMBER: 60/199, 384
;  PRIOR FILING DATE: 2000-04-25
;  NUMBER OF SEQ ID NOS: 2267
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 908
;   LENGTH: 302
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (262)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (279)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (294)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (295)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-11988-908


  Query Match             61.4%;  Score 205;  DB 1;  Length 302;
  Best Local Similarity   68.6%;  Pred. No. 2.5e-16;
  Matches   35;  Conservative    7;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      9 LLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
          ||||||  :||||  |||:|: ||||:| ||::|| |:|| | ||||||||
Db      2 LLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 52


RESULT   5
US-09-833-245-908
; Sequence 908, Application US/09833245
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Albumin Fusion Proteins
;  FILE REFERENCE: PF546PCT
;  CURRENT APPLICATION NUMBER: US/09/833,245
;  CURRENT FILING DATE:  2001-04-12
;  PRIOR APPLICATION NUMBER: 60/229, 358
;  PRIOR FILING DATE: 2000-04-12
;  PRIOR APPLICATION NUMBER: 60/256, 931
;  PRIOR FILING DATE: 2000-12-21
;  PRIOR APPLICATION NUMBER: 60/199, 384
;  PRIOR FILING DATE: 2000-04-25
;  NUMBER OF SEQ ID NOS: 2267
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 908
;   LENGTH: 302
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: SITE
;   LOCATION: (262)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (279)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (294)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (295)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-908


  Query Match           61.4%;  Score 205;  DB 5;  Length 302;
  Best Local Similarity  68.6%;  Pred. No. 2.5e-16;
  Matches   35;  Conservative    7;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      9 LLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
          ||||||  :||||  |||:|: ||||:| ||::|| |:|| | ||||||||
Db      2 LLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 52


RESULT   6
US-09-526-193A-37
; Sequence 37, Application US/09526193A
; GENERAL INFORMATION:
;  APPLICANT: Hayden, Michael R.
;  APPLICANT:  Brooks-Wilson, Angela R.
;  APPLICANT:  Pimstone, Simon N.
;  TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
;  TITLE OF INVENTION:  CHOLESTEROL LEVELS
;  FILE REFERENCE: 50110/002005
;  CURRENT APPLICATION NUMBER: US/09/526,193A
;  CURRENT FILING DATE:  2000-03-15
;  PRIOR APPLICATION NUMBER: 60/124,702
;  PRIOR FILING DATE: 1999-03-15
;  PRIOR APPLICATION NUMBER: 60/138,048
;  PRIOR FILING DATE: 1999-06-08
;  PRIOR APPLICATION NUMBER: 60/139,600
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: 60/151,977
;  PRIOR FILING DATE: 1999-09-01
;  NUMBER OF SEQ ID NOS: 287
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-526-193A-37


  Query Match           35.0%;  Score 117;  DB 5;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 4.9e-07;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     41 SVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||
Db      2 SVRLSYPPYEQHECHFPNKA 21


RESULT   7
US-60-248-505-1176
; Sequence 1176, Application US/60248505
; GENERAL INFORMATION:
;  APPLICANT: Beasley, Ellen
;  TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
;  TITLE OF INVENTION:  RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: cl000918
;  CURRENT APPLICATION NUMBER: US/60/248,505
;  CURRENT FILING DATE:  2000-11-15
;  NUMBER OF SEQ ID NOS: 1998
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1176
;   LENGTH: 774
;   TYPE: PRT
;   ORGANISM: Human
US-60-248-505-1176


  Query Match           18.7%;  Score 62.5;  DB 6;  Length 774;
  Best Local Similarity  22.8%;  Pred. No. 15;
  Matches   18;  Conservative    9;  Mismatches   21;  Indels   31;  Gaps    2;

Qy      8 RLLLWKNLTFRRRQTCQLLLEVAW----------------------------PLFIFLILIS 41
          :|::| :|||  :  || ||   |                            |:    |||
Db     53 KLVMWPHLTFVAQPNCQQLLATLWYDGFPGWRRKHWVVKLLTCMTIGFLFPMLSIAYLIS 112

Qy     42 VRLSY-----PPYEQHECH 55
            | :       |: :  ||
Db    113 PRSNLGLFIKKPFIKFICH 131


RESULT   8
PCT-US01-04098A-1326
; Sequence 1326, Application PC/TUS0104098A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-029
;  CURRENT APPLICATION NUMBER: PCT/US01/04098A
;  CURRENT FILING DATE:  2001-02-05
;  PRIOR APPLICATION NUMBER: Not Yet Assigned
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: 09/728,422
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 09/693,325
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/663,561
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 09/654,936
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 09/620,325
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/598,075
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 3960
;  SOFTWARE: Custom
; SEQ ID NO 1326
;   LENGTH: 551
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-04098A-1326


  Query Match           16.9%;  Score 56.5;  DB 1;  Length 551;
  Best Local Similarity  31.6%;  Pred. No. 59;
  Matches   18;  Conservative    8;  Mismatches   22;  Indels    9;  Gaps    2;

Qy      4 WPQLRLLLWKNLTFRRRQTCQLL-----LEVAWPLF----IFLILISVRLSYPPYEQ 51
          || || || :||  |  |  :       ||:|  |     : |: : :    || |:
Db    182 WPALRSLLHRNLVLRTHQPARYSLTPEGLELAQKLAESEGLSLLNVGIGPKEPPGEE 238


RESULT   9
PCT-US01-04098A-3294
; Sequence 3294, Application PC/TUS0104098A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
```

```
;  FILE REFERENCE: 21272-029
;  CURRENT APPLICATION NUMBER: PCT/US01/04098A
;  CURRENT FILING DATE:  2001-02-05
;  PRIOR APPLICATION NUMBER: Not Yet Assigned
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: 09/728,422
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 09/693,325
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/663,561
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 09/654,936
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 09/620,325
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/598,075
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 3960
;  SOFTWARE: Custom
; SEQ ID NO 3294
;   LENGTH: 595
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-04098A-3294


  Query Match             16.9%;  Score 56.5;  DB 1;  Length 595;
  Best Local Similarity   31.6%;  Pred. No. 62;
  Matches   18;  Conservative    8;  Mismatches   22;  Indels    9;  Gaps    2;

Qy      4 WPQLRLLLWKNLTFRRRQTCQLL-----LEVAWPLF----IFLILISVRLSYPPYEQ 51
          || || || :||  |  |  :       ||:|  |     : |: : :    || |:
Db    226 WPALRSLLHRNLVLRTHQPARYSLTPEGLELAQKLAESEGLSLLNVGIGPKEPPGEE 282


RESULT  10
PCT-US01-01310-71
; Sequence 71, Application PC/TUS0101310
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc., et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ03PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/01310
;  CURRENT FILING DATE:  2001-05-09
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-01310-71


  Query Match             16.8%;  Score 56;  DB 1;  Length 284;
  Best Local Similarity   24.2%;  Pred. No. 41;
  Matches   15;  Conservative   11;  Mismatches   10;  Indels   26;  Gaps    3;

Qy      8 RLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPY----------EQHECHFP 57
          |: ||          ||  :: :|| |       || :::||           |: :| |
Db    139 RMTLW---------TCAAVICMAWTL-------SVAMAFPPVFDVGTYKFIREEDQCIFE 182

Qy     58 NK 59
          ::
Db    183 HR 184


RESULT  11
PCT-US01-01310-77
; Sequence 77, Application PC/TUS0101310
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc., et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ03PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/01310
;  CURRENT FILING DATE:  2001-05-09
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77
;   LENGTH: 378
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-01310-77


  Query Match             16.8%;  Score 56;  DB 1;  Length 378;
  Best Local Similarity   24.2%;  Pred. No. 50;
  Matches   15;  Conservative   11;  Mismatches   10;  Indels   26;  Gaps    3;

Qy      8 RLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPY----------EQHECHFP 57
          |: ||          ||  :: :|| |       || :::||           |: :| |
Db    139 RMTLW---------TCAAVICMAWTL-------SVAMAFPPVFDVGTYKFIREEDQCIFE 182

Qy     58 NK 59
          ::
Db    183 HR 184


RESULT  12
PCT-US01-01310-102
; Sequence 102, Application PC/TUS0101310
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc., et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ03PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/01310
;  CURRENT FILING DATE:  2001-05-09
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (170)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01310-102


  Query Match             15.9%;  Score 53;  DB 1;  Length 183;
  Best Local Similarity   25.4%;  Pred. No. 65;
  Matches   15;  Conservative    9;  Mismatches    9;  Indels   26;  Gaps    3;

Qy      8 RLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPY----------EQHECHF 56
          |: ||          ||  :: :|| |       || :::||           |: :| |
Db    139 RMTLW---------TCAAVICMAWTL-------SVAMAFPPVFDVGTXKFIREEDKCIF 181


RESULT  13
PCT-US01-01332-806
; Sequence 806, Application PC/TUS0101332
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc., et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ02PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/01332
;  CURRENT FILING DATE:  2001-05-09
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1249
```

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 806
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (170)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01332-806


  Query Match            15.9%;  Score 53;  DB 1;  Length 183;
  Best Local Similarity  25.4%;  Pred. No. 65;
  Matches   15;  Conservative    9;  Mismatches    9;  Indels   26;  Gaps    3;

Qy       8 RLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPY----------EQHECHF 56
           |: ||         ||  :: :|| |       || :::||           |: :| |
Db     139 RMTLW---------TCAAVICMAWTL-------SVAMAFPPVFDVGTXKFIREEDKCIF 181


RESULT  14
US-09-383-745-1
; Sequence 1, Application US/09383745
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  TITLE OF INVENTION: 14926 Receptor, A Novel G-Protein Coupled Receptor
;  FILE REFERENCE: 035800/169197
;  CURRENT APPLICATION NUMBER: US/09/383,745
;  CURRENT FILING DATE:  1999-08-26
;  PRIOR APPLICATION NUMBER: 09/145,745
;  PRIOR FILING DATE: 1998-09-02
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 370
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-383-745-1


  Query Match            15.9%;  Score 53;  DB 5;  Length 370;
  Best Local Similarity  25.9%;  Pred. No. 1.1e+02;
  Matches   15;  Conservative   11;  Mismatches   12;  Indels   20;  Gaps    3;

Qy      13 KNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPY----------EQHECHFPNKA 60
           | |||    ||  :: : | |       || :::||           |: :| | :::
Db     131 KRLTF---WTCLAVICMVWTL-------SVAMAFPPVLDVGTYSFIREEDQCTFQHRS 178


RESULT  15
PCT-US01-01310-79
; Sequence 79, Application PC/TUS0101310
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc., et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ03PCT
;  CURRENT APPLICATION NUMBER: PCT/US01/01310
;  CURRENT FILING DATE:  2001-05-09
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
;   LENGTH: 379
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US01-01310-79


  Query Match            15.9%;  Score 53;  DB 1;  Length 379;
  Best Local Similarity  25.9%;  Pred. No. 1.1e+02;
  Matches   15;  Conservative   11;  Mismatches   12;  Indels   20;  Gaps    3;

Qy      13 KNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPY----------EQHECHFPNKA 60
           | |||    ||  :: : | |       || :::||           |: :| | :::
Db     140 KRLTF---WTCLAVICMVWTL-------SVAMAFPPVLDVGTYSFIREEDQCTFQHRS 187


Search completed: May 31, 2001, 13:08:13
Job time: 294 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 31, 2001, 12:20:09 ; Search time 92.94 Seconds
(without alignments)
103.844 Million cell updates/sec

Title: US-09-526-193A-1_COPY_1_60
Perfect score: 334
Sequence: 1 MACWPQLRLLLWKNLTFRRR.....SVRLSYPPYEQHECHFPNKA 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	334	100.0	162	23	US-60-169-629-486	Sequence 486, App
2	334	100.0	162	23	US-60-187-470-486	Sequence 486, App
3	334	100.0	2261	1	PCT-US00-06745-1	Sequence 1, Appli
4	334	100.0	2261	19	US-09-526-193-1	Sequence 1, Appli
5	334	100.0	2261	20	US-09-654-323-5	Sequence 5, Appli
6	230	68.9	101	23	US-60-206-111-183	Sequence 183, App
7	230	68.9	145	1	PCT-US00-30628A-175	Sequence 175, App
8	230	68.9	2083	23	US-60-230-445-1448	Sequence 1448, Ap
9	230	68.9	2180	23	US-60-221-839-5	Sequence 5, Appli
10	230	68.9	2180	23	US-60-230-445-2002	Sequence 2002, Ap
11	205	61.4	302	1	PCT-US00-30628A-101	Sequence 101, App
12	180	53.9	145	1	PCT-US00-30628A-178	Sequence 178, App
13	180	53.9	1446	23	US-60-258-275-311	Sequence 311, App
14	180	53.9	2235	14	US-09-032-438-6	Sequence 6, Appli
15	180	53.9	2273	14	US-09-032-438-3	Sequence 3, Appli
16	120	35.9	1875	23	US-60-230-445-1535	Sequence 1535, Ap
17	117	35.0	21	1	PCT-US00-06745-37	Sequence 37, Appl
18	117	35.0	21	19	US-09-526-193-37	Sequence 37, Appl
19	104	31.1	210	23	US-60-213-846-938	Sequence 938, App
20	104	31.1	2436	21	US-09-795-693-8	Sequence 8, Appli
21	104	31.1	2436	23	US-60-232-685-2	Sequence 2, Appli
22	103	30.8	2436	1	PCT-US00-40789-2	Sequence 2, Appli
23	86	25.7	1704	11	US-08-720-614-75	Sequence 75, Appl
24	86	25.7	2319	23	US-60-207-583-427	Sequence 427, App
25	86	25.7	2431	23	US-60-230-445-1964	Sequence 1964, Ap
26	86	25.7	3040	23	US-60-230-445-1210	Sequence 1210, Ap
27	78	23.4	1713	23	US-60-171-625-45	Sequence 45, Appl
28	78	23.4	1713	23	US-60-173-464-1728	Sequence 1728, Ap
29	78	23.4	1713	23	US-60-191-637-2064	Sequence 2064, Ap
30	78	23.4	1713	23	US-60-191-681-1647	Sequence 1647, Ap
31	78	23.4	1713	23	US-60-219-005-18	Sequence 18, Appl
32	70	21.0	100	23	US-60-147-499-4269	Sequence 4269, Ap
33	70	21.0	101	23	US-60-197-873-17098	Sequence 17098, A
34	70	21.0	1642	23	US-60-223-269-3	Sequence 3, Appli
35	63	18.9	63	1	PCT-US01-01354-13508	Sequence 13508, A
36	62.5	18.7	426	23	US-60-207-583-607	Sequence 607, App
37	62.5	18.7	426	23	US-60-230-445-1682	Sequence 1682, Ap
38	61.5	18.4	2271	23	US-60-173-464-27974	Sequence 27974, A
39	61.5	18.4	2272	23	US-60-191-637-36418	Sequence 36418, A
40	61.5	18.4	2272	23	US-60-191-681-28456	Sequence 28456, A
41	61	18.3	218	23	US-60-161-932-2601	Sequence 2601, Ap
42	60.5	18.1	464	18	US-09-450-969-4444	Sequence 4444, Ap
43	60.5	18.1	529	18	US-09-413-198-2403	Sequence 2403, Ap
44	59.5	17.8	181	18	US-09-417-507-38344	Sequence 38344, A
45	59.5	17.8	224	1	PCT-US93-09987-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-60-169-629-486
; Sequence 486, Application US/60169629
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: cDNAS for Secreted Proteins
; FILE REFERENCE: GENSET.071PRF
; CURRENT APPLICATION NUMBER: US/60/169,629
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 486
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -47..-1
US-60-169-629-486

Query Match 100.0%; Score 334; DB 23; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.8e-31;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
```

RESULT 2

```
US-60-187-470-486
; Sequence 486, Application US/60187470
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT: Bougueleret, L.
;  APPLICANT: Jobert, S.
;  TITLE OF INVENTION: cDNAS for Secreted Proteins
;  FILE REFERENCE: 78.US2.PRO
;  CURRENT APPLICATION NUMBER: US/60/187,470
;  CURRENT FILING DATE: 2000-03-06
;  NUMBER OF SEQ ID NOS: 715
;  SOFTWARE: Patent.pm
; SEQ ID NO 486
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -47..-1
US-60-187-470-486


  Query Match             100.0%;  Score 334;  DB 23;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-31;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   3
PCT-US00-06745-1
; Sequence 1, Application PC/TUS0006745
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  APPLICANT:  Xenon Bioresearch, Inc.
;  TITLE OF INVENTION:  METHODS AND REAGENTS FOR MODULATING
;  TITLE OF INVENTION:  CHOLESTEROL LEVELS
;  FILE REFERENCE:  50110/002WO5
;  CURRENT APPLICATION NUMBER: PCT/US00/06745
;  CURRENT FILING DATE:  2000-04-15
;  PRIOR APPLICATION NUMBER:  60/124,702
;  PRIOR FILING DATE:  1999-03-15
;  PRIOR APPLICATION NUMBER:  60/138,048
;  PRIOR FILING DATE:  1999-06-08
;  PRIOR APPLICATION NUMBER:  60/139,600
;  PRIOR FILING DATE:  1999-06-17
;  PRIOR APPLICATION NUMBER:  60/151,977
;  PRIOR FILING DATE:  1999-09-01
;  NUMBER OF SEQ ID NOS:  287
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2261
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US00-06745-1


  Query Match             100.0%;  Score 334;  DB 1;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-30;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   4
US-09-526-193-1
; Sequence 1, Application US/09526193
; GENERAL INFORMATION:
;  APPLICANT: Hayden, Michael R.
;  APPLICANT: Wilson, Angela R.
;  APPLICANT: Pimstone, Simon N.
;  TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
;  TITLE OF INVENTION: CHOLESTEROL LEVELS
;  FILE REFERENCE: 50110/002005
;  CURRENT APPLICATION NUMBER: US/09/526,193
;  CURRENT FILING DATE: 2000-03-15
;  EARLIER APPLICATION NUMBER: 60/124,702
;  EARLIER FILING DATE: 1999-03-15
;  EARLIER APPLICATION NUMBER: 60/138,048
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/139,600
;  EARLIER FILING DATE: 1999-06-17
;  EARLIER APPLICATION NUMBER: 60/151,977
;  EARLIER FILING DATE: 1999-09-01
;  NUMBER OF SEQ ID NOS: 287
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2261
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-526-193-1


  Query Match             100.0%;  Score 334;  DB 19;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-30;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   5
US-09-654-323-5
; Sequence 5, Application US/09654323
; GENERAL INFORMATION:
;  APPLICANT:  Hayden, Michael R.
;  APPLICANT:  Pimstone, Simon
;  APPLICANT:  Brooks-Wilson, Angela R.
;  APPLICANT:  Clee, Susanne  M.
;  TITLE OF INVENTION:  Compositions and Methods for Modulating
;  TITLE OF INVENTION:  HDL Cholesterol and Triglyceride Levels
;  FILE REFERENCE:  50110/004002
;  CURRENT APPLICATION NUMBER: US/09/654,323
;  CURRENT FILING DATE:  2000-09-01
;  PRIOR APPLICATION NUMBER:  US 60/124,702
;  PRIOR FILING DATE:  1999-03-15
;  PRIOR APPLICATION NUMBER:  US 60/138,048
;  PRIOR FILING DATE:  1999-06-08
;  PRIOR APPLICATION NUMBER:  US 60/139,600
;  PRIOR FILING DATE:  1999-06-17
;  PRIOR APPLICATION NUMBER:  US 60/151,977
;  PRIOR FILING DATE:  1999-09-01
;  PRIOR APPLICATION NUMBER:  US 09/526,193
;  PRIOR FILING DATE:  2000-03-15
;  PRIOR APPLICATION NUMBER:  US 60/213,958
;  PRIOR FILING DATE:  2000-06-23
;  NUMBER OF SEQ ID NOS:  256
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 2261
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-654-323-5


  Query Match             100.0%;  Score 334;  DB 20;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-30;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60


RESULT   6
US-60-206-111-183
; Sequence 183, Application US/60206111
; GENERAL INFORMATION:
;  APPLICANT:  Beasley, Ellen
;  TITLE OF INVENTION:  ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE:  CL000568
;  CURRENT APPLICATION NUMBER: US/60/206,111
;  CURRENT FILING DATE:  2000-05-22
;  NUMBER OF SEQ ID NOS:  234
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 183
;   LENGTH: 101
;   TYPE: PRT
;   ORGANISM: HUMAN
US-60-206-111-183


  Query Match             68.9%;  Score 230;  DB 23;  Length 101;
  Best Local Similarity   67.8%;  Pred. No. 2.6e-19;
  Matches   40;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
          || | || |||||| :|||| |||:|: ||||:| ||::|| |:|| | ||||||||
Db      1 MAFWTQLMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 59


RESULT   7
PCT-US00-30628A-175
; Sequence 175, Application PC/TUS0030628A
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: 28 Human Secreted Proteins
;  FILE REFERENCE: PS712PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/30628A
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 60/164,744
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/215,140
;  PRIOR FILING DATE: 2000-06-30
;  NUMBER OF SEQ ID NOS: 190
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US00-30628A-175


  Query Match             68.9%;  Score 230;  DB 1;  Length 145;
  Best Local Similarity   67.8%;  Pred. No. 3.5e-19;
  Matches   40;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
          || | || |||||| :|||| |||:|: ||||:| ||::|| |:|| | ||||||||
Db      1 MAFWTQLMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 59


RESULT   8
US-60-230-445-1448
; Sequence 1448, Application US/60230445
; GENERAL INFORMATION:
;  APPLICANT:  Beasley, Ellen
;  TITLE OF INVENTION:  ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE:  CL000765
;  CURRENT APPLICATION NUMBER: US/60/230,445
;  CURRENT FILING DATE:  2000-09-06
;  NUMBER OF SEQ ID NOS:  3051
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1448
;   LENGTH: 2083
;   TYPE: PRT
;   ORGANISM: HUMAN
US-60-230-445-1448


  Query Match             68.9%;  Score 230;  DB 23;  Length 2083;
  Best Local Similarity   67.8%;  Pred. No. 3.5e-18;
  Matches   40;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
          || | || |||||| :|||| |||:|: ||||:| ||::|| |:|| | ||||||||
Db      1 MAFWTQLMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 59


RESULT   9
US-60-221-839-5
; Sequence 5, Application US/60221839
; GENERAL INFORMATION:
;  APPLICANT:       Lal, Preeti
;  APPLICANT:  Yue, Henry
;  APPLICANT:  Walia, Narinder K.
;  APPLICANT:  Baughn, Mariah R.
;  APPLICANT:  Tribouley, Catherine M.
;  APPLICANT:  Yang, Junming
;  APPLICANT:  Thornton, Michael
;  APPLICANT:  Hafalia, April
;  APPLICANT:  Patterson, Chandra
;  APPLICANT:  Greene, Barrie D.
;  APPLICANT:  Yao, Monique G.
;  APPLICANT:  Raumann, Brigette E.
;  APPLICANT:  Gandhi, Ameena R.
;  APPLICANT:  Lu, Yan
;  APPLICANT:  Ding, Li
;  APPLICANT:  Tang, Y.Tom
;  APPLICANT:  Azimzai, Yalda
;  APPLICANT:  Burford, Neil
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Borowsky, Mark L.
;  APPLICANT:  Nguyen, Danniel B.
;  APPLICANT:  Khan, Farrah A.
;  APPLICANT:  Elliott, Vicki S.
;  APPLICANT:  Kearney, Liam
;  APPLICANT:  Lu, Dyung Aina M.
;  APPLICANT:  Thangavelu, Kavitha
;  APPLICANT:  Xu, Yuming
;  APPLICANT:  Sanjanwala, Madhu Sudan
;  APPLICANT:  Das, Debopriya
;  APPLICANT:  Policky, Jennifer L.
;  TITLE OF INVENTION:  TRANSPORTERS AND ION CHANNELS
;  FILE REFERENCE:  PI-0170 P
;  CURRENT APPLICATION NUMBER: US/60/221,839
;  CURRENT FILING DATE:  2000-07-28
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PERL Program
; SEQ ID NO 5
;   LENGTH: 2180
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7475603CD1
US-60-221-839-5


  Query Match             68.9%;  Score 230;  DB 23;  Length 2180;
  Best Local Similarity   67.8%;  Pred. No. 3.6e-18;
  Matches   40;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;
```

```
Qy       1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
           || | || |||||| :|||| |||:|: ||||:| ||::|| |:|| | ||||||||
Db      35 MAFWTQLMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 93


RESULT 10
US-60-230-445-2002
; Sequence 2002, Application US/60230445
; GENERAL INFORMATION:
;  APPLICANT:  Beasley, Ellen
;  TITLE OF INVENTION:  ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE:  CL000765
;  CURRENT APPLICATION NUMBER: US/60/230,445
;  CURRENT FILING DATE:  2000-09-06
;  NUMBER OF SEQ ID NOS:  3051
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2002
;   LENGTH: 2180
;   TYPE: PRT
;   ORGANISM: HUMAN
US-60-230-445-2002


  Query Match             68.9%;  Score 230;  DB 23;  Length 2180;
  Best Local Similarity   67.8%;  Pred. No. 3.6e-18;
  Matches   40;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
           || | || |||||| :|||| |||:|: ||||:| ||::|| |:|| | ||||||||
Db       1 MAFWTQLMLLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 59


RESULT 11
PCT-US00-30628A-101
; Sequence 101, Application PC/TUS0030628A
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: 28 Human Secreted Proteins
;  FILE REFERENCE: PS712PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/30628A
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 60/164,744
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/215,140
;  PRIOR FILING DATE: 2000-06-30
;  NUMBER OF SEQ ID NOS: 190
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
;   LENGTH: 302
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (262)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (279)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (294)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (295)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-30628A-101


  Query Match             61.4%;  Score 205;  DB 1;  Length 302;
  Best Local Similarity   68.6%;  Pred. No. 5.5e-16;
  Matches   35;  Conservative    7;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       9 LLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNK 59
           |||||| :|||| |||:|: ||||:| ||::|| |:|| | ||||||||
Db       2 LLLWKNFMYRRRQPVQLLVELLWPLFLFFILVAVRHSHPPLEHHECHFPNK 52


RESULT 12
PCT-US00-30628A-178
; Sequence 178, Application PC/TUS0030628A
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: 28 Human Secreted Proteins
;  FILE REFERENCE: PS712PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/30628A
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 60/164,744
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/215,140
;  PRIOR FILING DATE: 2000-06-30
;  NUMBER OF SEQ ID NOS: 190
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 178
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
PCT-US00-30628A-178


  Query Match             53.9%;  Score 180;  DB 1;  Length 145;
  Best Local Similarity   56.4%;  Pred. No. 2.5e-13;
  Matches   31;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       6 QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
           |::|||||| | |:||  : ::|: ||| :||:|| :| : | |  |||||||||
Db       6 QIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYSHHECHFPNKA 60


RESULT 13
US-60-258-275-311
; Sequence 311, Application US/60258275
; GENERAL INFORMATION:
;  APPLICANT: Beasley, Ellen
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001026-PROV
;  CURRENT APPLICATION NUMBER: US/60/258,275
;  CURRENT FILING DATE:  2000-12-27
;  NUMBER OF SEQ ID NOS: 717
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 311
;   LENGTH: 1446
;   TYPE: PRT
;   ORGANISM: HUMAN
US-60-258-275-311


  Query Match             53.9%;  Score 180;  DB 23;  Length 1446;
  Best Local Similarity   56.4%;  Pred. No. 1.8e-12;
  Matches   31;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       6 QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
           |::|||||| | |:||  : ::|: ||| :||:|| :| : | |  |||||||||
Db      32 QIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYSHHECHFPNKA 86


RESULT 14
US-09-032-438-6
; Sequence 6, Application US/09032438
;  GENERAL INFORMATION:
;    APPLICANT:  Allikmets, Rando, Anderson, Kent L., Dean, Michael, Leppert,
;    APPLICANT:  Mark, Lewis, Richard A., Li, Yixin, Lupski, James R., Nathans, Jerem
;    APPLICANT:  Amir, Shroyer, Noah F., Singh, Nanda, Smallwood, Philip, M., Sun, Hu
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
```

```
;    TITLE OF INVENTION:  FOR ATP-BINDING CASSETTE TRANSPORTER AND METHODS OF
;    TITLE OF INVENTION:  SCREENING FOR AGENTS THAT MODIFY ATP-BINDING CASSETTE
;    TITLE OF INVENTION:  TRANSPORTER
;    NUMBER OF SEQUENCES:  117
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & Norris
;      ADDRESSEE:  LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/032,438
;      FILING DATE:  27-FEB-1998
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  BYLR-0065
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2235 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-032-438-6


  Query Match             53.9%;  Score 180;  DB 14;  Length 2235;
  Best Local Similarity   56.4%;  Pred. No. 2.6e-12;
  Matches   31;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       6 QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
           |::|||||| | |:||  : ::|: ||| :||:|| :| : | |  |||||||||
Db       6 QIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYSHHECHFPNKA 60


RESULT  15
US-09-032-438-3
; Sequence 3, Application US/09032438
;  GENERAL INFORMATION:
;    APPLICANT:  Allikmets, Rando, Anderson, Kent L., Dean, Michael, Leppert,
;    APPLICANT:  Mark, Lewis, Richard A., Li, Yixin, Lupski, James R., Nathans, Jeremy,
;    APPLICANT:  Amir, Shroyer, Noah F., Singh, Nanda, Smallwood, Philip, M., Sun, Hui
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
;    TITLE OF INVENTION:  FOR ATP-BINDING CASSETTE TRANSPORTER AND METHODS OF
;    TITLE OF INVENTION:  SCREENING FOR AGENTS THAT MODIFY ATP-BINDING CASSETTE
;    TITLE OF INVENTION:  TRANSPORTER
;    NUMBER OF SEQUENCES:  117
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & Norris
;      ADDRESSEE:  LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/032,438
;      FILING DATE:  27-FEB-1998
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y
;      REGISTRATION NUMBER:  34,293
;      REFERENCE/DOCKET NUMBER:  BYLR-0065
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2273 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-032-438-3


  Query Match             53.9%;  Score 180;  DB 14;  Length 2273;
  Best Local Similarity   56.4%;  Pred. No. 2.6e-12;
  Matches   31;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       6 QLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
           |::|||||| | |:||  : ::|: ||| :||:|| :| : | |  |||||||||
Db       6 QIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYSHHECHFPNKA 60


Search completed: May 31, 2001, 13:07:08
Job time: 2819 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 31, 2001, 12:18:09 ; Search time 25.08 Seconds
(without alignments)
45.959 Million cell updates/sec

Title: US-09-526-193A-1_COPY_1_60
Perfect score: 334
Sequence: 1 MACWPQLRLLLWKNLTFRRR.....SVRLSYPPYEQHECHFPNKA 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	86	25.7	1704	3	US-08-762-500-75	Sequence 75, Appl
2	58.5	17.5	382	3	US-08-811-177A-2	Sequence 2, Appli
3	58	17.4	382	2	US-08-872-302-2	Sequence 2, Appli
4	56.5	16.9	1969	4	US-08-836-325-16	Sequence 16, Appl
5	56	16.8	921	1	US-07-872-644-39	Sequence 39, Appl
6	56	16.8	921	1	US-08-297-494-39	Sequence 39, Appl
7	56	16.8	921	1	US-08-297-510-39	Sequence 39, Appl
8	56	16.8	921	1	US-08-479-532-39	Sequence 39, Appl
9	56	16.8	921	1	US-08-455-526-39	Sequence 39, Appl
10	56	16.8	921	1	US-08-455-525-39	Sequence 39, Appl
11	56	16.8	921	3	US-09-139-491-39	Sequence 39, Appl
12	56	16.8	921	5	PCT-US92-03222-39	Sequence 39, Appl
13	56	16.8	941	1	US-07-872-644-45	Sequence 45, Appl
14	56	16.8	941	1	US-08-297-494-45	Sequence 45, Appl
15	56	16.8	941	1	US-08-297-510-45	Sequence 45, Appl
16	56	16.8	941	1	US-08-479-532-45	Sequence 45, Appl
17	56	16.8	941	1	US-08-455-526-45	Sequence 45, Appl
18	56	16.8	941	1	US-08-455-525-45	Sequence 45, Appl
19	56	16.8	941	3	US-09-139-491-45	Sequence 45, Appl
20	56	16.8	941	5	PCT-US92-03222-45	Sequence 45, Appl
21	56	16.8	942	1	US-07-872-644-43	Sequence 43, Appl
22	56	16.8	942	1	US-08-297-494-43	Sequence 43, Appl
23	56	16.8	942	1	US-08-297-510-43	Sequence 43, Appl
24	56	16.8	942	1	US-08-479-532-43	Sequence 43, Appl
25	56	16.8	942	1	US-08-455-526-43	Sequence 43, Appl
26	56	16.8	942	1	US-08-455-525-43	Sequence 43, Appl
27	56	16.8	942	3	US-09-139-491-43	Sequence 43, Appl
28	56	16.8	942	5	PCT-US92-03222-43	Sequence 43, Appl
29	56	16.8	1375	3	US-08-665-259-26	Sequence 26, Appl
30	56	16.8	1375	3	US-08-762-500-26	Sequence 26, Appl
31	55	16.5	232	3	US-09-024-020B-6	Sequence 6, Appli
32	55	16.5	383	1	US-08-314-596-41	Sequence 41, Appl
33	55	16.5	383	1	US-08-320-982-41	Sequence 41, Appl
34	55	16.5	383	3	US-08-819-037-41	Sequence 41, Appl
35	55	16.5	1976	3	US-09-024-020B-9	Sequence 9, Appli
36	55	16.5	1978	3	US-09-024-020B-3	Sequence 3, Appli
37	55	16.5	1988	3	US-09-024-020B-4	Sequence 4, Appli
38	54.5	16.3	1011	4	US-08-836-325-2	Sequence 2, Appli
39	54.5	16.3	1984	4	US-08-836-325-10	Sequence 10, Appl
40	54.5	16.3	1989	4	US-08-836-325-12	Sequence 12, Appl
41	53	15.9	370	3	US-09-251-373-2	Sequence 2, Appli
42	53	15.9	2005	4	US-08-836-325-7	Sequence 7, Appli
43	52.5	15.7	387	1	US-08-314-596-40	Sequence 40, Appl
44	52.5	15.7	387	1	US-08-314-596-42	Sequence 42, Appl
45	52.5	15.7	387	1	US-08-320-982-40	Sequence 40, Appl

ALIGNMENTS

```
RESULT 1
US-08-762-500-75
; Sequence 75, Application US/08762500
; Patent No. 6030806
; GENERAL INFORMATION:
;   APPLICANT: Landes, Gregory M.
;   APPLICANT: Burn, Timothy C.
;   APPLICANT: Connors, Timothy D.
;   APPLICANT: Dackowski, William R.
;   APPLICANT: Van Raay, Terence J.
;   APPLICANT: Klinger, Katherine W.
;   TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
;   TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
;   NUMBER OF SEQUENCES: 83
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: GENZYME CORPORATION
;     STREET: One Mountain Road
;     CITY: Framingham
;     STATE: Massachusetts
;     COUNTRY: United States of America
;     ZIP: 01701
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/762,500
;     FILING DATE: 09-DEC-1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/665,259
;     FILING DATE: 17-JUN-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US96/10469
;     FILING DATE: 17-JUN-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Dugan, Deborah A.
;     REGISTRATION NUMBER: 37,315
;     REFERENCE/DOCKET NUMBER: IG5-9.3
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (508) 872-8400
;     TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 75:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1704 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-762-500-75
```

```
  Query Match             25.7%;  Score 86;  DB 3;  Length 1704;
  Best Local Similarity   47.7%;  Pred. No. 0.0034;
  Matches   21;  Conservative    6;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRL 44
          ||   || |||||| | ::|:    :||:  ||    ||| :||
Db      1 MAVLRQLALLLWKNYTLQKRKVLVTVLELFLPLLFSGILIWLRL 44


RESULT   2
US-08-811-177A-2
; Sequence 2, Application US/08811177A
; Patent No. 6025172
;  GENERAL INFORMATION:
;    APPLICANT:  Dani, Maria
;    APPLICANT:  Catello, Sergio
;    TITLE OF INVENTION:  Isolation and Sequencing of
;    TITLE OF INVENTION:  the Hazel FAD2-N Gene
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, p.c.
;      STREET:  701-E 555 13th Street, N.W.
;      CITY:  Washington, DC
;      STATE:
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/811,177A
;      FILING DATE:  04-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G
;      REGISTRATION NUMBER:  35,377
;      REFERENCE/DOCKET NUMBER:  1729-331
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-783-6040
;      TELEFAX:  202-783-6031
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  382 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-811-177A-2


  Query Match             17.5%;  Score 58.5;  DB 3;  Length 382;
  Best Local Similarity   28.6%;  Pred. No. 3.1;
  Matches   10;  Conservative   10;  Mismatches   10;  Indels    5;  Gaps    1;

Qy     22 TCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHF 56
          |  : | : |||:     ::: :|  ||::  ||:
Db    180 TLLITLTLGWPLY-----LALNVSGRPYDRFACHY 209


RESULT   3
US-08-872-302-2
; Sequence 2, Application US/08872302
; Patent No. 5846784
;  GENERAL INFORMATION:
;    APPLICANT:  Hitz, William D
;    TITLE OF INVENTION:  Fatty Acid Modifying Enzymes From
;    TITLE OF INVENTION:  Developing Seeds of Vernonia galamenensis
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. duPont de Nemours and Co.
;      STREET:  1007 Market Street
;      CITY:  Wilmington
;      STATE:  Delaware
;      COUNTRY:  USA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/872,302
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Majarian, William R
;      REGISTRATION NUMBER:  P-41,173
;      REFERENCE/DOCKET NUMBER:  BB-1084
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-992-4926
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  382 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-872-302-2


  Query Match             17.4%;  Score 58;  DB 2;  Length 382;
  Best Local Similarity   36.4%;  Pred. No. 3.6;
  Matches   12;  Conservative    6;  Mismatches    9;  Indels    6;  Gaps    2;

Qy     27 LEVAWPLFIFLILISVRLSYPPYEQHECHF-PN 58
          | : |||::        :|  ||::  ||| ||
Db    185 LTLGWPLYLVF-----NVSGRPYDRFACHFSPN 212


RESULT   4
US-08-836-325-16
; Sequence 16, Application US/08836325
; Patent No. 6110672
;  GENERAL INFORMATION:
;    APPLICANT:  Mandel, Gail
;    APPLICANT:  Halegoua, Simon
;    APPLICANT:  Borden, Laurence A.
;    TITLE OF INVENTION:  Peripheral Nervous System Specific
;    TITLE OF INVENTION:  Sodium Channels, DNA Encoding Therefor, Crystallization,
;    TITLE OF INVENTION:  X-ray Diffraction, Computer Molecular Modeling, Rational
;    TITLE OF INVENTION:  Drug Design, Drug Screening, and Methods of Making and Usin
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C
;      STREET:  1100 New York Ave., N. W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/836,325
;      FILING DATE:  2-MAY-1997
;      CLASSIFICATION:  514
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/14251
;      FILING DATE:  02-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/482,401
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/334,029
;      FILING DATE:  02-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0917.0240002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1969 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-836-325-16


  Query Match             16.9%;  Score 56.5;  DB 4;  Length 1969;
  Best Local Similarity   22.4%;  Pred. No. 35;
  Matches   11;  Conservative   19;  Mismatches   18;  Indels    1;  Gaps    1;

Qy      10 LLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPN 58
           |  ::::   |::| :::|| |  :| :: :: : | :  |  :: :| |
Db    1190 LAFEDIYIERKKTIKIILEYADKIFTYIFILEMLLKWIAY-GYKTYFTN 1237


RESULT   5
US-07-872-644-39
; Sequence 39, Application US/07872644
; Patent No. 5389527
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/872,644
;      FILING DATE:  19920420
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5389527and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  921 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-872-644-39


  Query Match             16.8%;  Score 56;  DB 1;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
            |:     :|   | |:: |     || ||:|| ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT   6
US-08-297-494-39
; Sequence 39, Application US/08297494
; Patent No. 5580771
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/297,494
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5580771and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  921 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-297-494-39


  Query Match             16.8%;  Score 56;  DB 1;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
```

```
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:      :|    | |:: |     || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT   7
US-08-297-510-39
; Sequence 39, Application US/08297510
; Patent No. 5602019
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/297,510
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5602019and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  921 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-297-510-39


  Query Match             16.8%;  Score 56;  DB 1;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:      :|    | |:: |     || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT   8
US-08-479-532-39
; Sequence 39, Application US/08479532
; Patent No. 5776752
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,532
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/297,494
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5776752and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  921 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-479-532-39


  Query Match             16.8%;  Score 56;  DB 1;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:      :|    | |:: |     || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT   9
US-08-455-526-39
; Sequence 39, Application US/08455526
; Patent No. 5789553
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
```

```
;       COUNTRY:  USA
;       ZIP:  60603
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/455,526
;       FILING DATE:  31-MAY-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/297,494
;       FILING DATE:  29-AUG-1994
;       APPLICATION NUMBER:  US 07/688,356
;       FILING DATE:  04-APR-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5789553and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  27866/30822
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 346-5750
;       TELEFAX:  (312) 984-9740
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  39:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  921 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-455-526-39


  Query Match             16.8%;  Score 56;  DB 1;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:      :|    | |:: |     || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT  10
US-08-455-525-39
; Sequence 39, Application US/08455525
; Patent No. 5800987
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,525
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/297,494
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5800987and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  921 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-455-525-39


  Query Match             16.8%;  Score 56;  DB 1;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:      :|    | |:: |     || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT  11
US-09-139-491-39
; Sequence 39, Application US/09139491
; Patent No. 6015677
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/139,491
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,525
;      FILING DATE:  31-MAY-1995
;      APPLICATION NUMBER:  08/297,494
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6015677and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
```

```
;       TELEFAX:  (312) 984-9740
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  39:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  921 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-139-491-39


  Query Match             16.8%;  Score 56;  DB 3;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:     :|    | |:: |     || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT  12
PCT-US92-03222-39
; Sequence 39, Application PC/TUS9203222
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/03222
;      FILING DATE:  19920420
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Noland, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  921 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US92-03222-39


  Query Match             16.8%;  Score 56;  DB 5;  Length 921;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:     :|    | |:: |     || ||:||  ||  |  :|| |
Db     349 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 394


RESULT  13
US-07-872-644-45
; Sequence 45, Application US/07872644
; Patent No. 5389527
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/872,644
;      FILING DATE:  19920420
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5389527and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  941 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-872-644-45


  Query Match             16.8%;  Score 56;  DB 1;  Length 941;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:     :|    | |:: |     || ||:||  ||  |  :|| |
Db     369 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 414


RESULT  14
US-08-297-494-45
; Sequence 45, Application US/08297494
; Patent No. 5580771
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
```

```
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/297,494
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5580771and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  941 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-297-494-45


  Query Match             16.8%;  Score 56;  DB 1;  Length 941;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:     :|   | |:: |     || ||:||  ||  |  :|| |
Db     369 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 414


RESULT  15
US-08-297-510-45
; Sequence 45, Application US/08297510
; Patent No. 5602019
;  GENERAL INFORMATION:
;    APPLICANT:  Beavo, Joseph A.
;    APPLICANT:  Bentley, Kelley
;    APPLICANT:  Charbonneau, Harry
;    APPLICANT:  Sonnenburg, William K.
;    TITLE OF INVENTION:  DNA Encoding Mammalian
;    TITLE OF INVENTION:  Phosphodiesterases
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/297,510
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,356
;      FILING DATE:  04-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5602019and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30822
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  941 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-297-510-45


  Query Match             16.8%;  Score 56;  DB 1;  Length 941;
  Best Local Similarity   37.0%;  Pred. No. 18;
  Matches   17;  Conservative    6;  Mismatches   19;  Indels    4;  Gaps    1;

Qy       3 CWPQLRLLLWKNLTFRRRQ----TCQLLLEVAWPLFIFLILISVRL 44
           |:     :|   | |:: |     || ||:||  ||  |  :|| |
Db     369 CFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLL 414


Search completed: May 31, 2001, 13:04:32
Job time: 2783 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 31, 2001, 12:15:29 ; Search time 46.78 Seconds
(without alignments)
73.318 Million cell updates/sec

Title: US-09-526-193A-1_COPY_1_60
Perfect score: 334
Sequence: 1 MACWPQLRLLLWKNLTFRRR.....SVRLSYPPYEQHECHFPNKA 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS6/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	334	100.0	2143	21	B38108	Human ABC1 cholest
2	334	100.0	2259	21	B38107	Human ABC1 FHA-3 m
3	334	100.0	2260	21	B38106	Human ABC1 cholest
4	334	100.0	2261	21	B38082	Human ABC1 cholest
5	334	100.0	2261	21	B38104	Human ABC1 cholest
6	334	100.0	2261	21	B38105	Human ABC1 cholest
7	334	100.0	2261	21	B38109	Human ABC1 cholest
8	334	100.0	2261	21	B38110	Human ABC1 cholest
9	334	100.0	2261	21	B38111	Human ABC1 cholest
10	334	100.0	2261	21	B38112	Human ABC1 cholest
11	334	100.0	2261	21	B38113	Human ABC1 cholest
12	334	100.0	2261	21	B38114	Human ABC1 cholest
13	334	100.0	2261	21	B38115	Human ABC1 cholest
14	334	100.0	2261	21	B38116	Human ABC1 cholest
15	334	100.0	2261	21	B38117	Human ABC1 cholest
16	230	68.9	112	21	B41102	Human ORFX ORF866
17	180	53.9	2273	19	W70398	ATP binding casset
18	117	35.0	21	21	B38083	Human ABC1 antigen
19	86	25.7	1704	19	W46771	Amino acid sequenc
20	60.5	18.1	502	20	Y49625	Corn hexose carrie
21	59.5	17.8	224	15	R53701	Sequence of castor
22	58.5	17.5	382	18	W31740	Delta-12 desaturas
23	58.5	17.5	4473	17	R97244	Virulence gene clu
24	58	17.4	382	20	W83353	Vernonia galamenen
25	57.5	17.2	255	21	Y70405	Class 1 fatty acid
26	56.5	16.9	272	21	Y83086	F-box protein FBP-
27	56.5	16.9	1387	21	Y95441	Caenorhabditis ele
28	56.5	16.9	1977	17	R99641	Peripheral nervous
29	56	16.8	373	20	Y30534	A G protein-couple
30	56	16.8	373	20	Y30538	A G protein-couple
31	56	16.8	373	21	Y71300	Human orphan G pro
32	56	16.8	373	21	B02834	Human G protein co
33	56	16.8	373	21	Y32237	Human 7TM receptor
34	56	16.8	610	20	Y35481	Chlamydia pneumoni
35	56	16.8	921	16	R69727	Cyclic-GMP stimula
36	56	16.8	921	18	W18048	Cyclic-GMP-stimula
37	56	16.8	921	18	W11252	Clone p3CGS-5 cycl
38	56	16.8	921	19	W71224	cGS-PDE encoded by
39	56	16.8	921	19	W77040	Adrenal cortex Ca2
40	56	16.8	921	19	W60752	cGS-PDE isolated f
41	56	16.8	921	21	Y80984	Bovine adrenal cor
42	56	16.8	941	13	R28409	Human foetal cGS P
43	56	16.8	941	16	R69729	Cyclic-GMP stimula
44	56	16.8	941	18	W18050	Human cGS-PDE amin
45	56	16.8	941	18	W11253	pHcgs6n cyclic GMP

ALIGNMENTS

RESULT 1
B38108
ID B38108 standard; Protein; 2143 AA.
XX
AC B38108;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter FHA-1 mutant protein (R2144STOP).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.

PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR N-PSDB; C69389.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 shown on pages 152-157.
XX
SQ Sequence 2143 AA;

Query Match 100.0%; Score 334; DB 21; Length 2143;
Best Local Similarity 100.0%; Pred. No. 2.6e-37;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60

RESULT 2
B38107
ID B38107 standard; Protein; 2259 AA.
XX
AC B38107;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 FHA-3 mutant protein (delta-E1893, D1894).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR N-PSDB; C69388.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered

```
CC   risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 shown on pages 152-157.
XX
SQ   Sequence   2259 AA;


  Query Match             100.0%;  Score 334;  DB 21;  Length 2259;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT   3
B38106
ID   B38106 standard; Protein; 2260 AA.
XX
AC   B38106;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter FHA-1 mutant protein (delta-L693).
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW   mutein.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
DR   N-PSDB; C69387.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page -; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents a mutant human ABC1 cholesterol transporter
CC   associated with an altered cholesterol level and therefore an altered
CC   risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 shown on pages 152-157.
XX
SQ   Sequence   2260 AA;


  Query Match             100.0%;  Score 334;  DB 21;  Length 2260;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT   4
B38082
ID   B38082 standard; Protein; 2261 AA.
XX
AC   B38082;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
```

PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR N-PSDB; C69120.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Claim 5; Page 152-157; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents the human ABC1 cholesterol transporter.
XX
SQ Sequence 2261 AA;

Query Match 100.0%; Score 334; DB 21; Length 2261;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60

RESULT 5
B38104
ID B38104 standard; Protein; 2261 AA.
XX
AC B38104;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter TD-1 mutant protein (C1477R).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR N-PSDB; C69385.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 shown on pages 152-157.

XX
SQ Sequence 2261 AA;

Query Match 100.0%; Score 334; DB 21; Length 2261;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60

RESULT 6
B38105
ID B38105 standard; Protein; 2261 AA.
XX
AC B38105;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter TD-2 mutant protein (Q597R).
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
DR N-PSDB; C69386.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 shown on pages 152-157.
XX
SQ Sequence 2261 AA;

Query Match 100.0%; Score 334; DB 21; Length 2261;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60

RESULT 7
B38109
ID B38109 standard; Protein; 2261 AA.
XX
AC B38109;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter mutant, R219K.
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.

```
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page -; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents a mutant human ABC1 cholesterol transporter
CC   associated with an altered cholesterol level and therefore an altered
CC   risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 shown on pages 152-157.
XX
SQ   Sequence   2261 AA;


  Query Match             100.0%;  Score 334;  DB 21;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT   8
B38110
ID   B38110 standard; Protein; 2261 AA.
XX
AC   B38110;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter mutant, V399A.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW   mutein.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page -; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents a mutant human ABC1 cholesterol transporter
CC   associated with an altered cholesterol level and therefore an altered
CC   risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 shown on pages 152-157.
```

```
XX
SQ   Sequence   2261 AA;


  Query Match             100.0%;  Score 334;  DB 21;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT   9
B38111
ID   B38111 standard; Protein; 2261 AA.
XX
AC   B38111;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter mutant, V771M.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW   mutein.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page -; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents a mutant human ABC1 cholesterol transporter
CC   associated with an altered cholesterol level and therefore an altered
CC   risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 shown on pages 152-157.
XX
SQ   Sequence   2261 AA;


  Query Match             100.0%;  Score 334;  DB 21;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT  10
B38112
ID   B38112 standard; Protein; 2261 AA.
XX
AC   B38112;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter mutant, T774P.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW   mutein.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
```

PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 shown on pages 152-157.
XX
SQ Sequence 2261 AA;

Query Match 100.0%; Score 334; DB 21; Length 2261;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60

RESULT 11
B38113
ID B38113 standard; Protein; 2261 AA.
XX
AC B38113;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter mutant, K776N.
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 shown on pages 152-157.
XX

```
SQ   Sequence   2261 AA;

  Query Match           100.0%;  Score 334;  DB 21;  Length 2261;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT  12
B38114
ID   B38114 standard; Protein; 2261 AA.
XX
AC   B38114;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter mutant, E1172D.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW   mutein.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page -; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents a mutant human ABC1 cholesterol transporter
CC   associated with an altered cholesterol level and therefore an altered
CC   risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 shown on pages 152-157.
XX
SQ   Sequence   2261 AA;

  Query Match           100.0%;  Score 334;  DB 21;  Length 2261;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT  13
B38115
ID   B38115 standard; Protein; 2261 AA.
XX
AC   B38115;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter mutant, R1587K.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW   mutein.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
```

XX
DR WPI; 2000-587528/55.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 shown on pages 152-157.
XX
SQ Sequence 2261 AA;

Query Match 100.0%; Score 334; DB 21; Length 2261;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60

RESULT 14
B38116
ID B38116 standard; Protein; 2261 AA.
XX
AC B38116;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human ABC1 cholesterol transporter mutant, S1731C.
XX
KW Human ABC1 cholesterol transporter; chromosome 9q31;
KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW cardiovascular disease; coronary artery disease; coronary restenosis;
KW cerebrovascular disease; peripheral vascular disease;
KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
PN WO200055318-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-IB00532.
XX
PR 15-MAR-1999; 99US-0124702.
PR 08-JUN-1999; 99US-0138048.
PR 17-JUN-1999; 99US-0139600.
PR 01-SEP-1999; 99US-0151977.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
PA (XENO-) XENON BIORESEARCH INC.
XX
PI Hayden MR, Wilson AR, Pimstone SN;
XX
DR WPI; 2000-587528/55.
XX
PT New ABC1 polypeptide is useful for treating diseases associated with
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT disease and cancer -
XX
PS Examples; Page -; 229pp; English.
XX
CC The invention relates to the human ABC1 cholesterol transporter protein
CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC a member of the ATP-binding cassette (ABC transporter) superfamily of
CC proteins, and plays a crucial role in cholesterol transport, particularly
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC located on chromosome 9q31, and mutations in this gene are associated
CC with two genetic HDL (high density lipoprotein) deficiency disorders,
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC are distinguishable in that TD is an autosomal recessive disorder, while
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC cholesterol") in the blood correlate with a high risk of cardiovascular
CC disease, particularly coronary artery disease, but also cerebrovascular
CC disease, coronary restenosis, and peripheral vascular disease.
CC Conversely, a high level of HDL has protective effects against
CC cardiovascular disease. The invention provides genetic constructs and
CC transgenic cells and non-human animals comprising human ABC1 nucleic
CC acids, and methods of gene therapy for the treatment or prevention of
CC cardiovascular disease comprising the administration of an expression
CC vector encoding ABC1 or an active fragment thereof. The invention also
CC encompasses compounds which mimic ABC1 activity, compounds which
CC stimulate ABC1 expression and methods of screening for such compounds.
CC It further relates to methods for determining whether a patient has an
CC increased risk for cardiovascular disease due to polymorphisms in the
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC or prevent cardiovascular disease, especially coronary artery disease,
CC cerebrovascular disease, coronary restenosis or peripheral vascular
CC disease. They may also be used in the treatment of diseases associated
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC The invention specifically excludes proteins with the exact amino acid
CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC present sequence represents a mutant human ABC1 cholesterol transporter
CC associated with an altered cholesterol level and therefore an altered
CC risk of cardiovascular disease.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the native human ABC1 shown on pages 152-157.
XX
SQ Sequence 2261 AA;

```
  Query Match             100.0%;  Score 334;  DB 21;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


RESULT  15
B38117
ID   B38117 standard; Protein; 2261 AA.
XX
AC   B38117;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Human ABC1 cholesterol transporter mutant, I883M.
XX
KW   Human ABC1 cholesterol transporter; chromosome 9q31;
KW   ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;
KW   Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;
KW   cardiovascular disease; coronary artery disease; coronary restenosis;
KW   cerebrovascular disease; peripheral vascular disease;
KW   Alzheimer's disease; Niemann-Pick disease; Huntington's disease;
KW   X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;
KW   prognosis; prophylaxis; drug screening; transgenic animal; mutant;
KW   mutein.
XX
OS   Homo sapiens.
XX
PN   WO200055318-A2.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-IB00532.
XX
PR   15-MAR-1999;   99US-0124702.
PR   08-JUN-1999;   99US-0138048.
PR   17-JUN-1999;   99US-0139600.
PR   01-SEP-1999;   99US-0151977.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
PA   (XENO-) XENON BIORESEARCH INC.
XX
PI   Hayden MR,  Wilson AR,  Pimstone SN;
XX
DR   WPI; 2000-587528/55.
XX
PT   New ABC1 polypeptide is useful for treating diseases associated with
PT   ABC1 biological activity, e.g. Alzheimer's disease, Huntington's
PT   disease and cancer -
XX
PS   Examples; Page -; 229pp; English.
XX
CC   The invention relates to the human ABC1 cholesterol transporter protein
CC   (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is
CC   a member of the ATP-binding cassette (ABC transporter) superfamily of
CC   proteins, and plays a crucial role in cholesterol transport, particularly
CC   intracellular cholesterol trafficking in monocytes and fibroblasts, being
CC   involved in cholesterol efflux from the cell. The gene encoding ABC1 is
CC   located on chromosome 9q31, and mutations in this gene are associated
CC   with two genetic HDL (high density lipoprotein) deficiency disorders,
CC   Tangier disease (TD) and familial HDL deficiency (FHA). These diseases
CC   are distinguishable in that TD is an autosomal recessive disorder, while
CC   FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good
CC   cholesterol") in the blood correlate with a high risk of cardiovascular
CC   disease, particularly coronary artery disease, but also cerebrovascular
CC   disease, coronary restenosis, and peripheral vascular disease.
CC   Conversely, a high level of HDL has protective effects against
CC   cardiovascular disease. The invention provides genetic constructs and
CC   transgenic cells and non-human animals comprising human ABC1 nucleic
CC   acids, and methods of gene therapy for the treatment or prevention of
CC   cardiovascular disease comprising the administration of an expression
CC   vector encoding ABC1 or an active fragment thereof. The invention also
CC   encompasses compounds which mimic ABC1 activity, compounds which
CC   stimulate ABC1 expression and methods of screening for such compounds.
CC   It further relates to methods for determining whether a patient has an
CC   increased risk for cardiovascular disease due to polymorphisms in the
CC   ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat
CC   or prevent cardiovascular disease, especially coronary artery disease,
CC   cerebrovascular disease, coronary restenosis or peripheral vascular
CC   disease. They may also be used in the treatment of diseases associated
CC   with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick
CC   disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.
CC   The invention specifically excludes proteins with the exact amino acid
CC   sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic
CC   acid with the exact sequence as GenBank Accession No: AJ012376.1. The
CC   present sequence represents a mutant human ABC1 cholesterol transporter
CC   associated with an altered cholesterol level and therefore an altered
CC   risk of cardiovascular disease.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the native human ABC1 shown on pages 152-157.
XX
SQ   Sequence   2261 AA;


  Query Match             100.0%;  Score 334;  DB 21;  Length 2261;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-37;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 macwpqlrlllwknltfrrrqtcqlllevawplfiflilisvrlsyppyeqhechfpnka 60


Search completed: May 31, 2001, 13:03:56
Job time: 2907 sec
```